

How to talk about star wars

Mrs Margaret Thatcher's visit to Washington has clarified US intentions towards negotiations of ballistic missile defences. But the United States should be more forthcoming yet if it wishes.

TIME, perhaps only a few days, will tell whether the piece of paper Mrs Margaret Thatcher collected from President Ronald Reagan on 21 December will help smooth the way for the arms control negotiations that open next week at Geneva. Ever since President Reagan surprised the world with his announcement last March of a plan to win immunity from ballistic missiles carrying nuclear warheads, the Soviet government has been complaining bitterly about the threat that space will be the next big battlefield. Mr Mikhail Gorbuchov, on his otherwise successful visit to London two weeks ago, kept on at the same theme, and won cautious sympathy from Mrs Thatcher for his pains. What she has now done is to win from the United States administration an important clarification of the star wars plan, otherwise known as the Strategic Defense Initiative, or SDI. First, it is accepted, SDI is for the time being merely a research programme, and as such is not a violation of existing treaties (among which the ABM treaty, which limits anti-ballistic missile defences, is the most relevant). Second, the products of SDI (whatever they may be) will not be deployed without warning, and in particular without negotiation with the Soviet Union. That will ensure that Mrs Thatcher cannot be accused by her Soviet visitor of double-talk. Whether it will cut much ice at Geneva will depend on what the Soviet government makes of Washington's present account of its intentions.

There has been a sharp change of tenor in the past few months. In the first flush of star wars, the general assumption was that there had begun a huge new programme of military preparation whose only certain consequence would be to stimulate a similar range of developments in the Soviet Union. It is true, as President Reagan made quite clear, that nobody supposed two years ago that it would be possible to construct an effective defence against ballistic missiles overnight, with the result that the first vision of star wars foresaw a series of developments that would not be completed until the next century is well under way. But there have been many in the United States and elsewhere who have supposed that the momentum of SDI would continue from the first phase of introspection and research through to the deployment of weapons with increasing sophistication. If the US administration now says (as it does) that this was never the intention, it has only itself (or its president) to blame. For while Mr Reagan's original speech did make quite plain the speculative character of what was planned, it conspicuously failed to acknowledge the two chief causes of anxiety bound up with this adventure, fear that the intermediate stages in the building of a hypothetical perfect defence against ballistic missiles would dangerously upset the nuclear balance, and the suspicion that the United States did not care a fig for its obligations under the ABM treaty. That is why the understanding Mrs Thatcher has won is more than a mere token. It should help to make the whole star wars plan more accessible to rational discussion.

Mirage

Over nearly two years, there has been a string of objections to the whole concept of SDI, of which the most rudimentary has been that President Reagan's vision is an insubstantial mirage. This was the substance of the commentary on the proposals by the Office of Technology Assessment of the US Congress just over a year ago, as well as of more recent objections by Dr Sydney Drell's

committee, operating under the flag of the Union of Concerned Scientists in the United States. Such arguments turn on calculations of the numbers of space platforms that will be needed to keep all possible launching grounds for hostile missiles under surveillance, or of the amount of chemical or other fuel that would have to be carried into orbit so as to ensure that space-mounted defensive weapons could function. Inevitably, there has also been a running argument about the utility of a defensive weapons system in which each of several layers would inevitably be leaky, less than fully effective. What Washington in its present stance can fairly say is that these objections, which would be valid and even fatal if directed against the deployment of a star wars defence, do not apply to a research programme.

In reality, the programme now under way is a natural extension of what President Reagan was saying nearly two years ago, and soon afterwards made tangible by what the Pentagon's Fletcher panel reported in October 1983. The first line of defence in SDI is to be a system for attacking (and, with luck, destroying) hostile missiles when they are most easily identifiable, when their rockets are still burning in the lower stratosphere. Thereafter, in mid-course, detection is more difficult (because targets are less conspicuous) while destruction is more complicated (because each rocket will have multiplied into several warheads together with a cloud of decoys). For these two phases of interception, the technical hope seems to be some form of beam weapon that could deliver substantial amounts of energy over distances of the order of thousands of kilometres in times measured in seconds. (Interest seems now to centre on beams of neutral hydrogen atoms travelling with velocities of up to 0.9 of the velocity of light.) Later phases in a defence against ballistic missiles would be more conventional, relying on detection (and discrimination) from the ground, or within the atmosphere, and interception by ground-based rockets systems carrying conventional (not nuclear) warheads or even mechanical obstacles (such as the system of chains deployed in last year's successful US test of a missile interceptor). The crucial ingredient of any working system would be an automatic means of acquiring and manipulating data in such a way as to direct attacks at missiles, not decoys, that would outmatch in sophistication any electronic computing system yet built. Whatever the other uncertainties in the system, the feasibility of embodying in a parcel of electronics the equivalent of a whole crew of chiefs of staff, and their accompanying intelligence networks, must beggar the imagination of almost anybody.

Verification

So long as this remains a programme of research, it makes no sense to complain about it, or even to attempt to negotiate its abandonment. If arms control agreements are to be (as they should be) susceptible to verification, they can never include undertakings not to carry out research whose objectives can never be unambiguously defined. In a different field, if an agreement on chemical weapons were to include an interdiction of research, many ordinary laboratory experiments in organic synthesis would be disallowed. But is there not a danger that research, by acquiring momentum, will lead automatically to deployment? The implication of that question is a confession of despair in the



rationality of those who carry out research, even military research. If the critics of the feasibility of SDI are right, they will be fortified by events and star wars will never happen. Otherwise, the outcome will turn on negotiations. That, for the time being, should satisfy both the Soviet government and those who doubt the soundness of what the administration is about.

Three important questions remain, of which the chief is the relationship between SDI and the ABM treaty. The treaty (signed by the Soviet Union and the United States in 1972) was born of the conviction that the integrity of the retaliatory forces of the two superpowers would be the best guarantor of international security. The United States made this point explicit in a statement on 9 May 1972. The treaty itself requires that neither of the two superpowers should "develop, test or deploy" ABM systems, or components of them, which are "sea-based, air-based, space-based" or based on dry land and at the same time mobile. The purpose of these exclusions was to allow for the installation of two ABM systems in each country, an option partially taken up by Soviet Union (which has built a defensive system around Moscow) but eschewed by the United States (disappointed by the performance of its own system based on the Sprint rocket). On the face of things, last year's test of an anti-ballistic missile over the Pacific was already a violation of the treaty. The question when the SDI programme comes to seem a more deliberate violation would ordinarily occupy the lawyers for years to come were it not that the issue is too urgent to wait. The obvious difficulty is that even missile detection systems that could be intended as components of early warning systems, such as the satellite included in the next shuttle payload which is designed to monitor electromagnetic signals from eastern Asia, may now be thought to violate the treaty. Putting even simple particle accelerators into orbit is bound to give offence. At the very least, the United States had better look for a new understanding on the ABM treaty before it finds itself being accused of having made a mockery of the agreement. That, in present circumstances, is merely good international housekeeping.

The second issue, which will arise only if the first phase of SDI suggests that it would be feasible to build a defence of some kind against ballistic missiles, is whether it would ever be possible to construct such a system safely. There is much in what the US administration is now saying that even a less than perfect defence by each of four or five layers in a defensive system would make it impossible for a potential attacker to carry out with certainty a pre-emptive strike against its opponent's retaliatory force. Simple probability will ensure that. The calculation now is not (as in 1983) that the United States would be able to dispense with formal negotiations with the Soviet Union, secure in the knowledge that its deterrent force would always be safe, but that the installation of a ballistic missile defence would present opponents with the choice between a ruinously expensive programme of missile construction and hard-headed negotiations of the reduction of strategic missile forces. The snag, of course, is that for a period of perhaps two or three decades, the ballistic missile defence would be so far from perfect that pre-emptive strikes, either against the system itself or against the missiles it is intended to protect, would be more tempting than at present. And while each partner in the international superpower league may insist that its intentions are pacific, it must prudently calculate that its opponent will think the opposite, and will feel as threatened now by neutralizing of its own strategic forces as in 1972, when anti-ballistic missile systems were first reckoned to be paradoxically dangerous.

The third issue that stands out is pedestrian and financial. The first five years of SDI is reckoned to cost \$26,000 million, of which the US Congress has so far appropriated about five per cent. By the heady yardsticks of the US defence budget, this is not for the time being a large sum of money, but by other standards it is an enormous sum, especially when it is acknowledged that most of the money would be spent on the purchase of technical skills urgently needed in other fields. At a time when the US administration is hunting around for economies to make elsewhere in its use of public revenues so as to avoid the dangers of instability stemming from fiscal imbalance, it is an extravagance

to be avoided like the plague. So that too is another part of the case for asking that the United States should go further than the undertaking given to Mrs Thatcher, and that it should get next week's Geneva talks off in the right spirit not merely by volunteering a moratorium on anti-satellite weapons but by asking for a reinterpretation of the ABM treaty that will trade a licence to continue under the umbrella of SDI with a reaffirmation of the ABM treaty, which itself was always intended as a prelude to reductions of strategic arms. □

British broadcasting

The British government's perennial problem of financing the BBC cannot be shuffled off.

THERE are three things Britain does better than anywhere else: landscape gardening, military pageantry and television. Mrs Margaret Thatcher, the Prime Minister, would put the excellence of British television at risk by having the British Broadcasting Corporation (BBC) take advertising. She should not for the following reasons: (1) Even a little would spoil the pleasure of one of the few places on the globe where one can watch television without the annoying intrusion of advertising, like a little rubbish tip on a very beautiful mountain. (2) It would not solve the BBC's crisis, which is to have, in terms of annual revenue, something like half what commercial television enjoys (£600 million to £1,000 million).

The BBC itself has been arrogant in thinking it can escape real cuts in services while the rest of the British public sector cannot. It should jettison something, and be seen to do so. The best candidate is the popular music radio channel, Radio 1; pop music — ask any teenager — sounds better when interlaced with commercials. Second best (only because it would save merely £20 million a year) is local radio, always a mistaken venture into the fringes of broadcasting where a national service never need have gone. Third choice is to amalgamate two radio channels, say 2 and 4, or 3 and 4. Or the BBC could start later in the morning and stop earlier at night.

The BBC, which is not supported by the government but which enjoys a monopoly to sell licences to those who pick up electromagnetic signals at a price fixed by the government, is quite right to say two things. First, the secret is to broaden the base of the licence fee. (Why should hotels buy a single licence for 250 rooms?) The second is that the pensioners and others who really cannot afford the £65 a year the BBC now wants should be subsidized by the social services. The fact is that old people gain more enjoyment and benefit from television than much of the population. For those with a decent income, £65 is not too much. They should be prepared to pay 18 pence a day for what sustains them.

The real problem is to narrow the gap between the BBC and commercial or independent television (ITV). The way to do that, in Mrs Thatcher's mind, is to reduce ITV's income, undermining the old joke question of what is the difference between an oil tycoon and a videotape editor on London Weekend Television. (Answer: the tycoon does not have London weighting, the extra pay for working in London.) ITV companies pay a levy on two thirds of their profits after an initial levy-free slice. The government is considering switching the levy to a revenues base. But it has done this in the past and found it useless when the ITV companies had a bad year. What is needed is a new more flexible formula based partly on profits, partly on levy. The result would be to reduce the ITV's willingness to pay any price for peace among its trade unions on the grounds that what it pays comes straight off the levy. It is this that unfairly encourages a brain drain of talent from BBC to ITV at every level, from performer to make-up girl to cameraman.

When all these remedies are at hand, the BBC and the government should apply them. Just as British motorways are so strikingly free of billboards — a relief to the eye after the United States or even France — so an unblemished television landscape should not be surrendered lightly. □

Halley's comet

Closet US-Soviet collaboration in space

Washington

US research groups, working "under the table" for over a year, have designed and developed several instruments to be flown on two Soviet spacecraft headed for Halley's comet. The unusual collaboration, although unofficial, has involved extensive contacts between US and Soviet scientists, including a visit by a University of Chicago physicist to the Soviet Space Research Institute laboratory in Moscow, where he supervised the installation of his instruments.

An official agreement between the United States and the Soviet Union covering cooperation in space was allowed by the Reagan administration to lapse in 1982, after the imposition of martial law in Poland. Low-level contacts between US and Soviet government officials have continued, however, and the collaboration on the Halley's comet mission is said to have had the blessing of the US Departments of State and Defense. And the National Aeronautics and Space Administration (NASA), provided financial support to the University of Chicago group, which designed and built two comet-dust detectors.

The director of that group, John Simpson, had received an invitation from Academician R. Z. Sagdeev of the Space Research Institute in September 1983 to build the detectors for the Soviet mission. Earlier that month, at a scientific meeting in the Netherlands, Simpson had described his new detector design, which measures density of dust; Simpson was too late to have his instrument included in the Giotto spacecraft — the European Space Agency

(ESA) mission to Halley's comet.

Simpson's laboratory began construction of the new instruments in March 1984, delivering the first prototype in May. The two spacecraft were launched a week apart last month, and will be the first of the five missions planned to reach the comet.

The collaboration was announced only after the first launch, apparently for fear of second thoughts by the Soviet or US governments in the face of extensive publicity. The University of Chicago group used as a go-between the Max Planck Institute in Lindau, West Germany, where

several instruments were designed for these and other Soviet spacecraft. The Central Research Institute in Hungary was also credited by Simpson, for "facilitating the incorporation of the US experiments" into the Soviet spacecraft.

Other US groups are involved in the Soviet mission. John Hsieh, a physicist at the University of Arizona, designed a neutral mass spectrometer, built at Lindau, while scientists at Arizona and Michigan are participating in the analysis of imaging and plasma-physics data.

The comet-dust analyser will be the only entirely US-built experiment on any mission to Halley's comet. The spacecraft are to rendezvous with the comet in March 1986. The initial findings on dust density may be of particular importance in determining how closely the other spacecraft may safely approach the comet.

Stephen Budiansky

Research secrecy

US agencies hamstringing contracts

Washington

THE US government is increasingly using prepublication review of research results to restrict the free flow of academic information, according to a new report from Harvard University. The report finds that recent secrecy regulations "go far afield of any reasonable definition of national security", and suggests that prepublication review is being used by federal agencies that have nothing to do with national security to suppress unwelcome research findings.

A directive issued by President Reagan in March 1983 (National Security Decision Directive 84) required government employees to submit work for prepublication review as a condition for access to classified material. But, says the Harvard report, of federal agencies are increasingly including prepublication review clauses in university research contracts not involving classified

material. The National Institutes of Health, the Environmental Protection Agency and the Food and Drug Administration have all presented contracts containing such clauses. And although the restrictions are, in theory, aimed at halting the flow of valuable technology to the Soviet Union, some of the most restrictive clauses have appeared in contracts for social-science research.

The Harvard report notes that some agencies have begun to insist on the right to modify the scope of contracted research without the researcher's assent; others even insist that agency officials be allowed to participate directly in the research.

Another way in which the government agencies have been tightening restrictions on the free flow of information is in security classification. An executive order issued in 1982 (no. 12356) allows government officials to impose classification restrictions after a project has started; this, according to the report, has deterred academic researchers from taking on certain non-classified projects that they fear might be classified later on. The order also required officials to err on the side of stricter classification when in doubt.

Although the Department of Defense last year agreed not to use export control regulations to control fundamental research, it is not yet clear whether this principle applies to other federal agencies. A major policy statement on technical information restrictions is expected from the White House early this year.

Harvard denies that its report is a "call for action", although it is sure to become a focus for discussion in the academic community, especially as Harvard plans to update the document periodically. The report was prepared by John Shattuck, vice-president for government, community and public affairs.

Tim Beardsley

Press on pork-barrel laboratories

Washington

DR Frank Press, president of the National Academy of Sciences, has taken the unusual step of writing to all 1,428 members of the academy, urging them to "play a leadership role" in opposing so-called pork-barrel appropriations for scientific facilities. "None of these direct political approaches to secure funds from the research budget would have occurred if they were not initiated by members of our own community", the letter states.

In the past two years, half a dozen universities — several with the aid of a Washington lobbying firm — have taken their appeals for construction grants for new laboratories directly to influential congressmen, bypassing the usual peer-review process. The first and most notorious were Catholic University and Columbia Univer-

sity. All told, pork-barrel appropriations for scientific facilities come to nearly \$100 million over the past two years.

Most of these special appropriations have been taken out of the budget of the Department of Energy (DoE); there appears to be greater reluctance in Congress to earmark the budgets of the National Institutes of Health or the National Science Foundation, which award most of their funds through well-established procedures of peer review of investigator-initiated proposals. DoE has a more informal review process that relies on advisory panels to recommend priorities to the agency. None of the pork-barrel projects in question had been considered by these panels, nor had the appropriate congressional committees held hearings to look into them.

Stephen Budiansky

Australian technology

Button gains from Jones

Melbourne

LAST month's Australian federal elections have ended with a reduced majority for Mr R.J. (Bob) Hawke's Labor government and a portfolio reshuffle of considerable moment for Australia's fast-fading technology-based industries. With ballot-counting still incomplete, it is plain that a 2 per cent swing against the government has replaced the forecast Labor landslide.

With 34 Senate seats (against 33 for the conservative opposition parties), the government has lost its last chance to control the upper house, where the Australian Democrats now hold the balance of power with seven probable seats, the most recent won by Dr Norman Sanders, former US citizen, Korean War pilot, research geomorphologist, Tasmanian parliamentarian and crusading conservationist. The Nuclear Disarmament Party (NDP, see *Nature* 29 November, p.395), in the person of Ms Jo Vallentine, looks set to gain one seat, but shaven-headed lawyer/rock star Mr Peter Garrett will probably lose to the Democrats' deputy leader, Senator Mason.

Pre-election talk of a technological superministry has been borne out to some degree by the annexation of several of the technological development functions of Mr Barry Jones's Department of Science and Technology (DST) by the politically more powerful Department of Industry and Commerce. Under the leadership of the government Senate leader Senator John Button, the new ministry will be known as the Department of Industry, Technology and Commerce (DITC), while Mr Jones's reduced bailiwick becomes the Department of Science.

Precisely how many sections of the old DST will be handed over is not yet clear, but they are thought to include the Management and Investment Companies Scheme (MICS, see *Nature* 307, 584; 1984), the National Biotechnology Scheme, and elements of the DST Policy Division. As well as attempting to revitalize basic scientific research and continuing to boost the public's perception of it, Mr Jones will assist Senator Button in the House of Representatives.

The acquisition of the Australian Industrial Research and Development Incentive Scheme (AIRDIS) will allow Senator Button's expanded DITC to operate the new 150 per cent research and development tax write-off provision, which had been inserted into Labor's modest campaign platform before the election. Three major industrial companies have already indicated that they are considering research and development activities in Australia to take advantage of the next tax scheme.

Senator Button will never be the high-tech visionary Mr Jones is (but then, who

could be?). He is, however, much more of a political infighter, and while his stance in refusing increased protection to the depressed heavy-engineering industry engendered hostility among the unions, his sponsorship of the steel and automobile industry rationalization plans has earned him considerable respect in industry circles. The government's replacement of protectionist trade minister Mr Bowen by the former finance minister Mr Dawkins must also be interpreted as indicating the importance it places on efficient industry restructuring.

Jeffrey SellarMedical research

Too little cash

THE British Medical Research Council (MRC) is still not sure how much it will have to spend in 1985-86, but one thing is already clear — there will not be enough. The removal of £3 million from the last-minute addition to the science budget (see *Nature* 312, 582; 1984) has made the future look even more bleak. The council plans to restore £1 million to the budgets of its research units (cut last autumn by just over £2 million), but otherwise fears there will be a further deterioration in its capacity to support applications for long-term (five years or more) research grants, perhaps by 25 per cent. The result of that could be that only two out of five first-class applications would succeed. MRC will also be more selective in its support for research units, aiming at the survival of the best at the expense of the less good.

Sir James Gowans, secretary of MRC, delivering this sombre message on the eve of the recent holiday, also called into question government policy on the science budget and in particular the British government's assumption that the Department of Education and Science must be solely responsible for finding the cash with which to support the research councils. The result, this year, had been that the department had had to look for extra science money by cutting other educational services. Gowans thinks the government should shoulder collective responsibility for science.

Without quite knowing where the funds will come from, MRC will nevertheless go ahead with four new projects on which it has set its heart, and which are:

- A unit for the application of molecular biology to medicine (at the John Radcliffe Hospital, Oxford).
- A new programme in neurobiology.
- Expansion of medical imaging techniques (Hammersmith Hospital, London).
- A centre for collaborative research with industry (Mill Hill, London).

Maxine Clarke

European research

Commission looks ahead

Brussels

ALTHOUGH the funds allocated to the European Communities' research and development programmes on 19 December fall well below the Commission's original proposals, there is now reason to hope for a substantial increase two years from now, when research ministers will select programmes with particular promise.

Last month's meeting agreed on a budget of 1,225 million European Currency Units (ECU, £2,040 million) for the four or five-year research programmes, some of which have been in limbo for more than a year. More than half the budget will go on thermonuclear fusion (690 million ECU). Radiation protection and waste management gets 120 million ECU, biotechnology 55 million ECU, the stimulation programme (by which the communities support national initiatives) 60 million ecu, the "Brite" programme for industrial research 125 million ECU and non-nuclear energy 175 million ECU.

The hope of extra funds is linked with the ministers' decision that two-thirds of last month's allocation should be spent within two years, whereupon there is to be a thorough review of the scientific value and Community interest in the different programmes. If everything prospers, the result may be that total spending over the next five years will not fall far short of the Commission's original proposals.

The Esprit programme in information technology, with a budget separately agreed, is to move into a new phase this year. The research meeting on 19 December agreed that new proposals would have to be more detailed than in the past, with more precise timetables.

The outgoing commissioner for research, Etienne Davignon, was mildly optimistic when he announced the outcome of the research ministers' meeting. National governments had made substantial compromises in reaching agreement on the new budget, he said. And research is no longer a second-string in Europe, where "fundamental research is the best in the world".

Davignon's concern, as always, is with commercial competition in technology; he urged that European companies should pay more attention to the need for self-sufficiency. He had no doubt that the Common Market could be made to function more efficiently, but not necessarily in time to keep up with the United States and Japan. And Davignon was also glum because the proposed increase of the Commission's budget, to be secured by an increased rate of value added tax of 1.4 per cent, had not yet been agreed by national parliaments but had "already been spent" in Brussels.

Anna Lubinska

Polish agriculture

Church and State at odds

POLAND'S controversial "Agricultural Foundation", established three months ago under the patronage of the Roman Catholic Church to assist the recovery of private agriculture, is now working out a programme of ten pilot projects, ranging from milk production to tractor repairs, and from water development to agricultural field services. But there are growing doubts about the foundation's effectiveness, depending as it does both on a massive injection of Western capital and on good relations between Church and State.

Uniquely in the socialist bloc, some 75 per cent of Poland's agricultural land is in private hands, in the form of peasant small-holdings, most of which are far smaller than the 15-20 hectares that the UN Food and Agriculture Organization considers the smallest viable size in Northern Europe. Moreover, almost three decades of under-investment and an artificial price structure that favoured the state sector have resulted in the gradual run-down of agriculture, although in 1980 the constraints on individual farmers were relaxed, and legislation introduced guaranteeing them security of tenure, thereby increasing their incentive to improve their holdings.

At the same time, several rescue operations for agriculture were proposed, including a project by the US Rockefeller Foundation that would have included significant support for the state sector. In the event, it was the Church's proposal, aiding the private sector only, which was adopted. This foundation hopes to raise

some \$1,800 million of external aid over five years; it will purchase external supplies with hard currency raised in the United States and Western Europe, and sell them to farmers at prevailing market prices, partly on a deferred-payment basis. These sales will eventually establish a zloty fund to be used to finance such in-country investments as artesian wells, and also to establish a social infrastructure in rural areas to help halt the drift of young people away from the land.

As yet there is no guarantee that the Agricultural Foundation will in fact receive sufficient backing from the West, and there are fears that the foundation will provoke much argument between Church and State for very little return. Although the land belongs to the farmers, water resources, for example, belong to the state. And although there is a free market in many agricultural products, there is no free market in grain, all of which must be sold to the state at fixed prices. Moreover, the foundation plans not only to provide farm machinery, but to establish a service infrastructure for the machinery.

Scepticism about the foundation comes also from agricultural academies and colleges — largely through ignorance of the foundation's plans. This is unfortunate, since once the foundation becomes operative, as now seems likely, extension courses to teach the farmers to make the best use of their resources and some kind of consultancy service will undoubtedly be necessary.

Vera Rich

Neutron source

A shoestring operation

BRITAIN'S spallation neutron source (SNS), which will soon be the world's most intense source of low-energy neutrons, produced its first test beams of neutrons last month at one four-thousandth of the design intensity. Full operation is due to start in April, but, says Dr Geoff Manning, director of the Rutherford-Appleton Laboratory where SNS is based, "at present we don't have enough money to run next year".

Running costs will be £10 million, he estimates, compared with a capital cost of £50 million of new money. But there is only £8 million in the kitty. The difference is equal to the expected electricity bill for operating SNS in 1985.

Nevertheless SNS will run — but in somewhat straitened circumstances. The machine and its four existing instruments, which worked perfectly in the test run, are already two years late as a result of funding delays, and is now in serious danger of under-exploitation. SNS is designed to run 20 instruments on 18 beam-lines, and to provide facilities for "several hundred" UK university and industrial scientists. But exactly when that maximum will be reached is unclear. The Science and Engineering Research Council has been actively seeking foreign participation and support, and has so far succeeded only with India, whose Bhabha Institute in Bombay has provided one of the four existing instruments.

There had been British hopes that West Germany would help, but faced with no British money for HERA (a proton-electron collider now under construction in Hamburg), nor for the European Synchrotron Radiation Source, nor any British commitment to a next-generation neutron spallation source planned in Germany for the 1990s, it is perhaps not surprising that the German government has been lukewarm in contributing to SNS. There will be one German experiment at SNS — a neutrino detector for experiments on masses and mixing of neutrino types.

SNS's main role, as was evidenced by the test experiments last week, involves neutrons as diffractive probes of atomic, molecular and material structures of liquid and solid-state matter. In this the use of neutrons is similar to that of X rays, with the exception that the energy of a neutron of a given (de Broglie) wavelength is much lower than that of an X ray of the same wavelength — thus allowing greater intensities without damage to the sample. Neutrons interact principally with nuclei, rather than the electronic structure of a material, thus providing complementary information to X-ray (and synchrotron radiation) investigations.

Robert Walgate

United States and UNESCO

Washington

THE new year starts with still no clear indications of what institutional arrangements will be made to allow the United States continued participation in the scientific programmes run by the United Nations Educational, Scientific and Cultural Organization (UNESCO). US withdrawal from the organization took effect from 1 January. A budget proposal for programmes to replace those of UNESCO has been sent by the State Department to the White House Office of Management and Budget, and details of the President's proposals will be released together with general budget figures later this month. These programmes will not, however, start until October 1985, and some extra stop-gap support may be needed until then. Furthermore, in spite of the State Department's "commitment in principle" to diverting the annual contribution saved from UNESCO into development, it now seems unlikely that the full \$47 million per year saved will be spent in this way.

A report on UNESCO's science programmes by the US National Research

Council has suggested that specific support through Funds-in-Trust could be appropriate for funding many programmes, while organizations such as the International Council of Scientific Unions might represent US interests in others. But even if UNESCO's science programmes are treated well in the budget for fiscal year 1986, international organizations may want some guarantees that US funds they administer will not suddenly dry up.

The US National Commission on UNESCO recently passed a resolution expressing "regret" at the withdrawal decision and urged the State Department to restore the commission's full-time staff to monitor further reforms at UNESCO. Gregory Newell, assistant secretary of international organization affairs at the State Department, acknowledged that the United States might consider rejoining the organization if its complaints were met, but hinted that the United States would consider pulling out of other international organizations that were damaging US interests or were not meeting their objectives.

Tim Beardsley

European pollution

Lead-free fuel on sale by 1989

Brussels

ALTHOUGH European Community environment ministers meeting in Brussels on 6 December agreed to introduce lead-free petrol in the Community from 1 October 1989, the decision to phase out lead from petrol may cause friction between countries keen to introduce controls and those less concerned about lead pollution. The directive seeks to avoid confrontation with West Germany, which said in September that it would start selling lead-free petrol from 1986.

At the end of a marathon session, ministers had agreed to make a single grade of "super-petrol" widely available in the Community on a mandatory basis from 1989, while at the same time leaving member states free to introduce the lead-free variety any time between now and then. The European Parliament's environment committee had recommended the introduction of lead-free petrol by 1986.

West Germany also plans to speed up the phasing out of normal leaded petrol, but a unilateral move will require prior consent of the European Commission and the other member states.

Although lead-free petrol will be on sale from 1989, the agreement reached by ministers does not include any incentives to encourage motorists to use lead-free petrol, in spite of recommendations by consumer groups in a Community report published in April 1984.

The only "technical" constraint to use lead-free petrol would be if cars were to be fitted with catalytic converters in order to conform with stricter vehicle emission standards. But ministers made little progress on the Commission's proposal to reduce vehicle emissions (carbon monoxide, nitrous oxides and hydrocarbons) to US standards and a decision on the matter has been postponed until the next Environment Council, scheduled for the end of January 1985.

In the meantime, proposals will be prepared for emission standards to be laid down according to the different categories of cars. It is likely that they will suggest the strictest standards for 2-litre models and over, which represent 10 per cent of cars on European roads, while cars under 1.3 litres might escape controls altogether. The idea of different categories does not, however, appeal to the West Germans, who are sticking to their decision to make anti-pollution devices obligatory on a national scale in 1988 and 1989.

The West German decision means that three-way or oxidation catalytic converters will soon be essential to all cars sold there. But countries that produce mainly small and medium-sized cars, such as France and Italy, are reticent in view of the high cost involved, preferring to wait for lean-burn engine technology which could make its

debut in Europe in two years' time.

A more immediate solution to curb emissions is to introduce stricter speed limits. The proposal put forward by French environment minister Huguette Bouchardeau, however, was not welcomed by West Germany, where speed on the autobahns is unrestricted.

On the other hand, pressure from the French was largely responsible for the watering down of the commitments made in June to limit the lead content of leaded petrol (which will be sold in parallel during the transition period for cars built before 1989) to 0.15 grams per litre. The French complained that the reduction from the present content of 0.40 grams per litre would cost them FF1,500 million. Under the new directive, this will still not be obligatory. Only West Germany and Denmark apply the stricter lead content, while the United Kingdom and the Netherlands plan to do so in 1986.

Cuban science

Soviet plan for joint technology

SOVIET scientific cooperation with Cuba will be increased under a new long-term economic and technical agreement concluded in November. The new programme foresees joint efforts in accelerating the development of Cuban science and technology, building up the material and technical research base and improving the qualifications of Cuban scientific cadres.

Arrangements will be worked out for the transfer of scientific and technical documentation to Cuba, and an educational aid programme will cover not only the training of Cuban specialists in technical schools and higher education institutes (presumably in both Cuba and the Soviet Union), but also "study attachments" for Cuban technologists, engineers and managers at Soviet enterprises.

The new cooperation agreement is intended to complete the construction of the "material and technical base of socialism" by the end of the century. It concentrates on nine main spheres: agriculture and food processing, fuel and energy, metallurgy, machine building, electronic and electrical engineering, chemical and light industry, medical industry, construction materials and the building industry, and transport. Science and technology, formally the tenth point of the plan, seems to be seen solely in terms of the other nine.

Although the agreement is intended to strengthen the Cuban economy, there is no intention that Cuba should be self-sufficient. On the contrary, the Havana conference of Comecon leaders last summer took the idea of intra-Comecon cooperation a stage further from the existing bilateral cooperation projects.

As far as other aspects of air pollution control that were on the agenda are concerned, ministers adopted a directive setting air quality limit values for nitrogen dioxide in the air of 200 micrograms per cubic metre of air and guide values of 50 micrograms per cubic metre. The main sources of nitrogen dioxide are in fact car exhausts and factory chimneys. The standards have been strongly criticized by both the European Parliament and the European Environmental Bureau for being far too high to produce any significant reduction in the damage done to man and the environment. No progress was made on the directive limiting air pollution from large combustion plants, which would cost the United Kingdom £4,000 million, according to the UK Central Electricity Generating Board.

Meanwhile, research is to continue. Ministers agreed to grant 4 million ECU for the first year of a five-year programme to collect information and to look at different methods of measuring, among other things, environmental degradation due to acid precipitation.

Anna Lubinska

From now on, the intention is that the five-year plans of all member countries will be coordinated, leading, theoretically, to more specialization and division of labour between them.

The new agreement specifically stresses the need to "help Cuba develop the production of specific types of output — primarily in the processing sectors of industry — that will ensure its broader participation in specialization and production-sharing arrangements within Comecon. Products specifically mentioned for export are sugar, fresh and processed citrus fruit, tobacco, possible new citrus products (vitamin C, lysine, citric acid) and medicinal plants for the pharmaceutical industry. There are less specific references to product-sharing and specialization in the electronics industry, a mutual exchange of chemical raw materials and finished products, the export of "sewn and other light industry" goods as part of a compensation deal, and, somewhat remarkably, the possible production of unspecified medical and laboratory apparatus for export to the Soviet Union.

Cuba still depends considerably on Soviet expertise and specialist advice. During an impromptu press conference last month, Dr Fidel Castro paid a glowing tribute to Soviet assistance which had provided Cuba not only with the model plant he was visiting but with oil refineries, harbour facilities and a nuclear power programme. Nevertheless, it appears that Cuba does not intend to rely solely on Soviet aid, but is prepared to shop around for expert advice within the bloc.

Vera Rich

Engineering

Better planning at NSF?

Washington

A RADICAL reorganization of the engineering directorate of the National Science Foundation (NSF) was blessed last week by the directorate's advisory council. The changes are being pushed by the foundation's new director for engineering, Nam P. Suh, and are aimed at providing "more coherent" planning of engineering science.

Suh argues that NSF should take more of the initiative in engineering research planning, identifying and building up areas where the basic science is not in place for technological development as well as building on existing strengths. To this end, Suh proposes to sweep away the existing divisions based on traditional subject categories (such as chemical, mechanical and civil engineering) and replace them with three basic engineering science divisions and divisions for "emerging engineering systems" (for example biotechnology process engineering), for "critical engineering systems" (projects in which public institutions may take a leading role, such as hazardous waste disposal and earthquake engineering) and for engineering where the science base is inadequate.

Suh denies that his aim is to seek funds for engineering research at the expense of fundamental science, and does not see a need for a major increase in federal funding for engineering research in general. The reorganization of the engineering directorate will not simply be a paper-pushing

exercise, he says, because the new teams will be "sending out the right signals to the research community". The aim is to create a "new culture" in which scientists are more free to step over traditional subject boundaries.

The other component of Suh's efforts to revitalize engineering is the proposed network of engineering research centres to be based at universities. The idea is that each centre should deal with a different end-use technology, such as computer-aided design or data and communications, providing links with industry and interdisciplinary research experience for students. The centres are intended to provide experience with industrial problems for up to 10 per cent of students of the host institutions (though some admit to doubts about whether that target will prove feasible).

NSF's budget for fiscal year 1985 includes a mere \$10 million for the engineering research centres, and with a proposed budget of \$2.5-\$5 million per year for each centre it is by no means clear how far NSF will get towards its target of 25 centres. The deadline for proposals was 1 October, by which time no fewer than 142 bids had been received from 107 different engineering schools. NSF is adopting a pragmatic approach, waiting to see what the President's budget proposal will contain in January before committing itself to a definite number. Site visits are now being planned.

Tim Beardsley

New institute tackles information technology

THE Turing Institute, a self-styled "centre of excellence in advanced information technology" in Glasgow, Scotland, was formally opened last month by Mr Geoffrey Pattie, UK Minister of State for Industry and Information Technology. With the help of start-up funds from the Scottish Development Agency, and with working capital provided by subscribing industrial companies, the institute promises elaborate services, including intensive training, discounted use of commercial and non-commercial software, library and information services, a "confidential" quarterly newsletter and consultancy.

Perhaps the most ambitious departure, however, is the Journeyman training scheme, simultaneously unveiled, which has been inaugurated by the Alvey Intelligent Knowledge Based Systems (IKBS) Directorate. From now on, it is hoped that the group of industrial affiliates that funds the institute by private subscription will send employees for six-month secondment programmes in logic programming, expert system building, applied expert systems and robotics.

Four centres were originally considered

for the scheme, including Malvern and Edinburgh. But only the department of computing at Imperial College London, will be joining Turing in the first phase of Journeyman next month. Two people are enrolled for the first course, although four will be the normal number once the project gets off the ground. If the scheme is successful, it may then be extended.

The Department of Trade and Industry (DTI) contributes the lion's share of Alvey's budget, about £250 million to industry's £100 million, and a government spokesman says that this is ample proof of DTI's "eagerness to encourage this kind of research". It is debatable, however, whether private institutions such as the Turing Institute and restricted retraining initiatives such as Journeyman are seriously analogous, as the department claims, to the substantial investment of artificial intelligence research and development in West Germany, Japan and the United States. But these provisions for accelerated "technology transfer" may mean that DTI is not content to rest on its laurels and the United Kingdom's traditional excellence in the field of machine intelligence.

Hugh Barnes

University of Wales

College merger engineered

THE governing bodies of University College, Cardiff (UCC) and of the University of Wales Institute of Science and Technology (UWIST) announced last month that they have decided to combine to form a new college within the University of Wales. Similar proposals in 1981 came to nothing, but this latest resolution came about after the University Grants Committee (UGC) received separate applications from both institutions for funds to develop separate engineering faculties. UGC favoured the establishment of a single integrated faculty to provide a strong engineering base for South Wales and a joint working party of the two colleges is now preparing a feasibility study on its behalf. It is hoped that a £7 million development of a site at Newport Road will provide much-needed and purpose-built space, offering the new faculty improved facilities for teaching and research.

UGC has given its assurance that "the grant provided to the new merged institutions would not be below the total of what would otherwise be available to the two separate institutions". This assurance, however, refers only to the first year of its existence, at the end of which resources and numbers will go into the melting-pot. The working party is confident that no compulsory redundancies in the teaching or technical staff will result from amalgamation. About 1,200 students will be enrolled at the new college, slightly more than the present aggregate. Indeed, the belief that a certain size is a critical requirement of success is behind much of the present planning — the new college will be among the six largest of its kind in Britain.

Professor Denis Towill, a member of the working party, has been involved in discussions since a merger was first suggested. He anticipates considerable opportunities for running otherwise uneconomic courses. "The new faculty will allow us to build up the traditional skills of South Wales engineering, those relevant to the steel industry, civil engineering, mineral exploitation, metallurgy, energy studies and electrical power engineering." Increased space, and a broader catchment area extending to Swansea on one side and Cheltenham on the other, Professor Towill claims, will bolster work in new technology, in microelectronics, digital computers and robotics.

The working party is due to report on its consultations with the architects early this month. If the plans go ahead, construction will inevitably be delayed until 1987. It is hoped, however, that the integrated faculty of engineering will be able to move into its new premises by 1990. Centralization of the other departments will take place over the following five years.

Hugh Barnes

Agricultural R & D

SIR — A recent publication¹ of the Organization for Economic Cooperation and Development (OECD) included a comment that “in the United States, Germany and France government spending on agricultural research and development is around one per cent of value added in the industry. The percentage . . . rises to over two per cent in the United Kingdom”. The same point has been taken up in the latest *Annual Review of Government Funded R & D*².

agricultural research and development but it can be argued that fishery research should not be. What is labelled as “agronomic” research in Table 1 derives from the collection of data on financial support for institutions rather than subjects; research from general university funds is shown under a separate category (promotion of knowledge) but this is then subdivided by subject area, one of which is agronomic research. Because, in West Germany especially, much agricultural

Table 1 Expenditure on agricultural research and development in European Community countries (1981) (£ million)

Item	Category	Ireland	Denmark	Belgium	Netherlands	France	UK	Italy	West Germany
1	Agricultural productivity and technology	11	13	15	48	143	129	34	78
2	Fishing and fishery products	2	(1)	1	1	13	7	1	6
3	Net agricultural productivity and technology	9	12	14	47	130	122	33	72
4	Food, drink and tobacco	2	(1)	2	(4)	18	3	5	10
5	Agronomic research	1	(4)	7	13	4	41	28	89
6	Bioscience	(3)	(4)	19	(11)	58	3	20	61
7	Total agricultural R & D	15	21	42	75	210	169	86	232

Values in parentheses inferred by interpolation on a £ per head basis.

A histogram is shown for France, Germany, Italy, Japan, the United Kingdom and the United States and it is suggested “that the UK spends, in GDP [Gross Domestic Product] terms, more than the other countries on agricultural R & D”. Combined with an earlier statement that, even if such indicators do not of themselves provide a basis for judgement, they do “indicate where questions over the balance of expenditure might be asked”, it would seem that publicly-funded agricultural research and development in the United Kingdom may now face yet another round of severe cuts.

How valid is this interpretation of the OECD indicators? An analysis based on Eurostat data³ for countries of the European Community suggests a rather different view. Essentially, the Eurostat data derive from the same primary sources as the OECD indicators and can be reconciled with them. The categorizations they use, however, enable a broad summary of agricultural research and development to be made (Table 1).

The OECD data are confined to item 1, “agricultural productivity and technology”. A more realistic estimate emerges, however, if, after eliminating a minor amount of fishery research and development, the national efforts on food, “agronomic research” and “bioscience” (which must include biotechnology) are included. The reasons for doing so are as follows.

Food research is normally reckoned as part of the remit of publicly-funded

research is carried out in universities, this effort fails to be captured by the OECD indicators. As is evident from Table 1, the effect is considerably to underestimate West German agricultural research and development.

There is a similar problem with “bioscience”, in that much UK non-medical work in this area is carried out in institutes under the general label of items 1 or 5. The UK bioscience effort in Table 1 is thus anomalously low, as will be evident to those familiar with this field. In a preliminary analysis, a temporary expedient is to count “bioscience”, for all countries, as agricultural research and development.

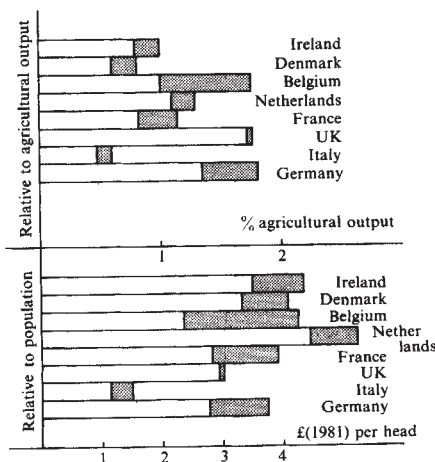


Fig. 1 Agricultural research and development for European Community countries relative to agricultural output and to population. The stippled area refers to bioscience.

The ambiguities inherent in the OECD indicators are illustrated by the transfer of funds in the “customer–contractor” arrangements. These were recorded as a rapid increase in expenditure on “agricultural productivity and technology”, a change which OECD itself pointed out was “more apparent than real”.

Relative to agricultural output, the UK level of agricultural research and development does not now appear greatly in excess of the levels in other countries (Fig. 1). (Italy is substantially below the others and does not provide an appropriate basis for comparison with the United Kingdom.) Moreover, relative to population (which could be taken as a proxy variable for national food consumption), the UK level of agricultural research and development is actually lower than in other Community countries. If substantial reductions were made beyond those already planned, the United Kingdom would begin to fall markedly behind the most important Community countries.

The whole subject needs wider and deeper analysis; meanwhile, the danger of putting too much reliance on empirical indicators is evident. In the end, agricultural research and development stands or falls by what it offers the nation; at a time when other countries are rapidly developing new technology based on biological resources (that is, agricultural research and development in the widest sense of the term⁴), it is disappointing that the *Annual Review* has failed to capture this spirit of enterprise.

W.S. WISE

*Agricultural and Food Research Council,
160 Great Portland Street,
London W1, UK*

1. *OECD Science and Technology Indicators: Resources devoted to R & D* (OECD, Paris, 1984).
2. *Cabinet Office. Annual Review of Government Funded R & D 1984* (HMSO, 1984).
3. *Government Financing of Research and Development 75–1982; Basic Statistics of the Community* (Eurostat, Luxembourg, 1983).
4. *The Widening Scope of Agricultural Research: Constraints and Challenges, 5th Working Conference of Directors of Agricultural Research* (OECD, Paris, 1984).

Scientists' oaths?

SIR — Jonathon Howard in his letter (*Nature* 8 November, p.96) is very flattering to the medical profession, implying a superior code of ethics.

While not denying this, I would point out that it is far from routine for newly qualified physicians and surgeons in the United Kingdom to be required to swear either the Hippocratic oath or have the Geneva modernization. It would appear that the expectations of our peers are adequate to dictate our standards.

Why should the same not be true for scientists?

CHARLES CLAOUÉ

*University of Bristol,
Department of Ophthalmology,
Bristol Eye Hospital,
Lower Maudlin Street,
Bristol BS1 2LX, UK*

Squarks and strings more real

Particle physics is again in turmoil. Experimentalists believe they may have new phenomena to describe. Theorists may already have an accommodating theory.

THESE are heady days for particle physics. Hardly a week after Dr Carlo Rubbia of CERN had been in Stockholm to collect his Nobel prize for discovering the W^+ and Z^0 particles, which mediate the weak nuclear force, he paid a flying visit to a meeting at the Rutherford Laboratory near Oxford to describe preliminary evidence that the same UA1 experiment may have revealed a counterpart of ordinary matter called "supermatter", in which "matter" (such as electrons) and "radiation" (such as photons) exchange roles.

Many theorists would dearly like to believe Rubbia's evidence. For example, John Ellis, also from CERN (the European Organization for Nuclear Research near Geneva), where Rubbia's UA1 experiment is taking new data on high energy proton-antiproton collisions, has risked a bottle of wine in a wager with US theorist Sydney Drell that supermatter will be discovered by next July. Rubbia himself was last week more cautious. Other mysterious events at the CERN collider have sunk under the flood of new data taken in the past three months, he explained. But certain "monojet" (now "bijet") events, with large amounts of missing energy, persist, cannot be explained as experimental artefacts but can be interpreted as evidence for the creation and decay of superparticles. The UA1 detectors around the proton-antiproton collider at CERN have just come to the end of a run, but it is clear from what Rubbia said two weeks ago that there is still a great deal to be quarried from the data already collected.

Rubbia hopes to be able to make a more confident statement on these events early in the year. Whether his group has truly found supermatter or not, what has the theoreticians agog is that it is already predicted by theories based on the idea of "supersymmetry", which endows every fundamental entity with a counterpart with contrasting quantum statistics. Familiarly, electrons with Fermi statistics obey Pauli's exclusion principle, making chemistry possible. Photons, on the other hand, have Bose statistics, which allows many photons to exist in the same state, and makes intense light beams and electromagnetic fields possible. But selectrons (selectrons) would have Bose statistics, while superphotons (photinos) would be fermions. Quarks (fermions) would be partnered by squarks (bosons), gluons (bosons) by gluinos (fermions), and so on.

Since supermatter seems not to exist at

low energy, theorists have hitherto assumed that the supermatter counterparts of ordinary particles have large masses. But theory has been able to put only vague limits on the masses of superparticles. One interpretation of Rubbia's events, compatible with these limits, is that he is creating 40-GeV squarks, paired with 40-GeV antisquarks, each of which decays into an ordinary quark (or antiquark) and a photino. The photinos interact only weakly with ordinary matter, and carry off the missing energy which is now thought to characterize the special UA1 events. The quark and antiquark would each make "jets" of hadronic matter.

Ellis had already predicted that, if squarks have a sufficiently low mass to be produced at UA1, the detector should see pairs of jets plus missing energy. His difficulty had been that Rubbia's early data had shown only single jets. Now, however, Rubbia's group is finding bijets — news that delighted Ellis, who heard it for the first time at the Rutherford meeting. Moreover, the UA1 group can now interpret the early detection of monojets as an artificial experimental bias against bijets. For example, many of the early "monojets" can be seen, on closer inspection, to have had other jets associated with them which were eliminated by a computer code deliberately designed to ignore jets below 12 GeV total energy.

As if the potential discovery of supermatter were not enough, Michael Green of Queen Mary College, London, introduced at the meeting a theory he has developed with John Schwarz of the California Institute of Technology that describes particles as massless relativistic "superstrings". Superstrings are entities in ten dimensions (nine space-like, one time-like) which are expected to behave like ordinary particles when the ten dimensions are collapsed to four, but whose characteristic properties derive from the six hidden dimensions.

For the past several years, theorists have been toying — and more than that — with particle theories written in more than four dimensions. One fashion, for example, has been a sequence of theories in eleven dimensions, in which the extra seven dimensions are supposed at the outset to be closed on themselves with a certain characteristic scale. The properties of particles in the real world can then be recovered as the bubble of non-existent space shrinks to nothing, like the smile on

the face of a Cheshire cat. The advantage is that manipulation in many dimensions may be easier than in a mere four.

The theory has already chalked up conspicuous successes, notably the accolade from Professor Edward Witten of Princeton that the theory is "stunning". One of its results is a simple relation between the gravitational constant and the gauge coupling constant (the "charge" which describes the fundamental strength of electromagnetic, strong and weak forces) of the form

$$\kappa = \text{constant} \cdot g^2 \cdot T$$

where κ is the Newtonian gravitational constant, g the gauge coupling and T (the tension in the fundamental string) is the only adjustable constant in the theory. This link is related to the simple fact that in this theory nongravitational forces are caused by the exchange of open-ended strings, and gravitational forces by the exchanges of closed loops. The existence of one kind of force implies the existence of the other.

Superstring theory is supersymmetric (so there would be no conflict with the discovery of supermatter), but the real reasons it has caught attention are its links with gravity, its apparent consistency with quantum mechanics and its apparent near-uniqueness.

Green and Schwarz's superstrings in a ten-dimensional space-time have the further advantage that they are inherently chiral, which means that they can describe parity-breaking, an essential for any theory meant to accommodate the weak interaction, as a truly unified theory of the forces must.

Here lies one of the "miracles" of the theory, as Green describes them. Existing grand unified theories are plagued by "anomalies" in charge conservation when they are quantized; the theories are constructed to conserve charge, but when certain simple interactions of particles are calculated, it turns out that charge conservation is violated, in most chiral versions of the theories, unless certain unnatural coincidences are assumed to occur. But in Green and Schwarz's superstring theory, the anomalies cancel each other naturally; moreover, this cancellation involves the exchange of loops, which is another way of saying that gravity is intimately involved. The hint, therefore, is that the long-sought quantum theory of gravity may also be hidden in the theory. It is no wonder that strings are suddenly so fashionable.

Robert Walgate

Ecology

How different are Australian ecosystems and ecologists?

from Mark Westoby

IN THE years between sabbaticals, the thinking of Australian ecologists can become developed — fixed, even — to the point where their ideas cannot be driven extinct just because they are inconsistent with what the fashionable majority in the northern hemisphere thinks. That is one possible reason why there have been so many influential Australian ecologists. But another possible reason is that Australian ecosystems are themselves different; to have experienced them is to know that the textbooks are wrong. Last year's symposium of the Ecological Society of Australia* was on that theme.

A central question in ecology is whether the general properties of assemblages of organisms are determined by the present-day environment, or whether they vary idiosyncratically depending on accidents of the biota's evolutionary history. Can ecologists understand ecosystems, or do we have to hand the job over to palaeontologists? Because Australia's biota has such a separate evolutionary history from the rest of the world, comparisons between Australia and elsewhere are natural experiments on this issue.

Many contributors at the meeting concluded that Australian assemblages are different — but that the differences arise from under-appreciated special properties of the Australian environment, not from evolutionary history. For example P.S. Lake (Monash University) pointed out that the eucalypt leaf-fall, which drives the food-chains of forest streams, arrives during summer while the water is warm, instead of being concentrated in autumn. A.V. Milewski and R.M. Cowling (Western Australian Institute of Technology) compared sclerophyll scrublands in western Australia and South Africa, looking at vegetation on several soil types within each continent. Features such as pollination and seed-dispersal spectra, as well as plant growth forms and leaf shapes, were influenced by soil type. Therefore comparisons between continents ideally need to be on precisely matched soils — but different continents might not contain precisely matched environments. R.J. Whelan (University of Wollongong) contrasted regeneration after fire in Florida and western Australia. In western Australia, many species have their dormancy broken by fire, and early regeneration is a flush of seedlings. In Florida, early regeneration is all vegetative, leading to flowering, production of non-dormant

seeds, and a flush of seedlings in the second year. Whelan argued that reliance on seeds, which are made germinable by fire, is only possible when a reliable rainfall season follows the fire season, as in western Australia.

Conclusions of this type are congenial for Australian ecologists: they mean that more research is needed in Australia. Nevertheless some ecologists were willing to emphasize how similar assemblages on different continents could be once one looked behind the taxonomy. B. Rice (unaffiliated) summarized patterns of local richness in plant species. The same vegetation types that are rich in species on one continent can be poor on another. Available data do not show any clear difference in species richness between different continents within a vegetation type. Present-day environment therefore seems to be a much more important influence than evolutionary history on local richness in plant species. Several contributors described cases where they believed evolutionary history had made a fundamental difference to Australian ecosystems. For example, P.C. Heyligers (CSIRO) and J.D. Sauer (University of California, Los Angeles) independently proposed that Australia lacked an effective sand-binding plant on

foredunes, with far-reaching consequences for landforms and vegetation in coastal areas. H.A. Ford (University of New England) suggested that bird pollination was common in Australia in compensation for the lack of social bees.

Methodology loomed large. It had, for example, been thought that insect herbivores take a larger proportion of leaf production in Australian forests than elsewhere. C.D. Ohmart (CSIRO), for eucalypt forests, and M. Lowman (University of New England, Armidale), for rain forests, argued that this was an artefact of using different or poor methods. A.J. Underwood and P.G. Fairweather (Sydney University), drawing on case histories from the study of intertidal communities, illustrated how different conclusions can flow from differences in study design or in the attitudes of investigators.

An unusually wide range of organisms and environments were considered at the meeting. This made it possible to go beyond a simple 'yes' or 'no' to the question of whether ecosystem properties are convergent. It seems clear that many ecosystem properties do converge, provided the environments are similar. But some quite subtle differences between environments can be important, and we are beginning to learn what differences to look for. And just occasionally, as in the case of plants that bind sand into foredunes, the presence or absence of an organism with a particular way of life can result in fundamental changes for the rest of the ecosystem. □

Mark Westoby is in the School of Biological Sciences at Macquarie University, North Ryde, NSW 2113, Australia.

Evolutionary biology

Group selection and the sex ratio

from Paul H. Harvey, Linda Partridge and Len Nunney

A DEEPER understanding of evolutionary processes often comes from looking at the same problem in more than one way¹. One problem that has benefited from this approach is that of explaining the female-biased sex ratio often found in hymenopterans (bee, ant, wasp) and mite populations^{2,3}.

In 1967, W.D. Hamilton pointed out that an unequal allocation of parental resources to sons and daughters should evolve under certain population structures². He produced a simple model to illustrate the process. Consider a population consisting wholly of fertilized females that come together in groups to lay their eggs. The young from the several clutches laid together mate randomly among themselves before the males die and the females disperse to form the new population in the next generation. Under those circumstances females investing either equally in sons and daughters or more in sons than in daughters would leave

fewer offspring than those producing an appropriately female-biased sex ratio.

Hamilton² suggested that a female-biased sex ratio was favoured because of competition for mates among siblings of the same sex — 'local mate competition'. In practice, sons are more likely to compete for mates than are daughters, because the reproductive success of a daughter is often limited by the number of eggs she can produce, whereas that of a son is limited by the number of mates available to him. If the sons of a single female compete with each other for mating opportunities then, for example, an isolated female need only produce a single son, provided that he is able to fertilize all of his sisters. The production of additional sons is wasted because they do not produce additional grandchildren, whereas additional daughters do, because production of grandchildren is essentially limited by eggs and not sperm. The expectation of a female-biased sex ratio holds in Hamilton's model even when more than

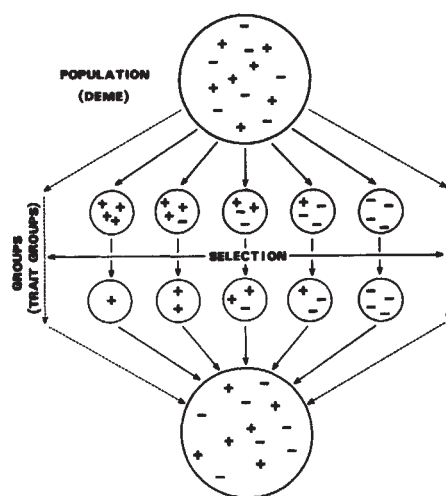
*'Are Australian Ecosystems Different?', Sydney, 28–29 August 1984. The proceedings will be published by the Society in early 1985 and distributed by Blackwell Scientific.

one female contributes sons and daughters to each mating group.

Thereafter Taylor⁴ and more recently Grafen⁵ have argued that not one but two processes are involved in producing the female-biased sex ratio in Hamilton's model. First, the number of offspring produced per son decreases with the number of sons in the brood. Second, females that produce more daughters can increase the reproductive success of their sons because, for example, the sons can mate with their additional sisters.

Modifications of Hamilton's original model, where either sib-mating is prevented or the number of offspring produced per son does not increase with additional sons, provide support for the view that there are indeed two processes at work^{6,7,8}. Furthermore, Maynard Smith⁸ has developed a model in which both processes are prevented. If females lay all of their daughters in one patch but each of their sons in a separate patch, the patch structure of the model is maintained, but the evolutionarily stable sex ratio is not biased towards sons or daughters.

Between them, these models, seem to illustrate a complete explanation of female-biased sex ratios in Hamilton's original model. Colwell, however, has argued that group selection is necessary to account for the biased sex ratio⁹. Group selection models have played an important role in evolutionary biology¹⁰ following Wynne-Edwards's suggestion that population densities of many species are held in check by the voluntary foregoing of breeding¹¹. This is done for the good of the local population (the group), preventing their overcropping of renewable resources. Wynne-Edwards's type of argument was countered by modelling subdivided populations containing both altruistic and selfish individuals^{12,13}.



Population structure envisaged under models of intrademic group selection. A global population (the deme) divides into groups (trait groups) where selection acts before the global population is reconstituted¹³. Trait groups can be short-lived or they can last for many generations⁸. Thereby one allele (-) of a gene can be selected against in mixed groups but nonetheless increase in frequency across the population compared with another allele (+).

Selfish individuals produce more offspring than altruists, but to the detriment of their group which either goes extinct or has reduced productivity. Because they are selected against within a mixed group but are favoured by selection among groups, genes for altruism can only increase in frequency by group selection. So long as there is at least limited migration of individuals between groups, individual selection within groups is a much stronger evolutionary force than inter-group selection¹⁴.

A slightly different type of population

Type of female	H		H		H		F		F		F	
No. of offspring	3 ♀	1 ♂	3 ♀	1 ♂	3 ♀	1 ♂	2 ♀	2 ♂	2 ♀	2 ♂	2 ♀	2 ♂
Grandchildren per child	4	12	4	12	4	6.7	4	6.7	4	4	4	4
Total grandchildren	12 + 12		12 + 12		12 + 6.7		8 + 13.3		8 + 8		8 + 8	

Fertilized females from a large population randomly assort into groups of two. Each has 4 offspring. 'Hamilton' (H) females produce three female and one male, while 'Fisher' (F) females produce two of each sex. The offspring mate within the group. Each female offspring will have four children. The reproductive success of male offspring is determined by how many mates are available within the group and how many other males are competing for those mates. (From ref. 1).

structure was described in a series of papers published in the 1970s (ref. 15). The global population (deme) divides into groups (trait groups), within which selection operates and from which individuals disperse to reconstitute the global population (see figure). Intrademic group selection will result in an increase in the frequency of a gene *a* if selection in mixed groups favours the alternative allele *b*; nevertheless, *a* increases in frequency in the global population. The two most likely reasons for this are trait groups arising by non-random sampling from the global population or non-additive fitness interactions among individuals with different genotypes¹.

Colwell's claim that group selection is responsible for female-biased sex ratios is based on two facts. First, females producing young with the evolutionarily stable ('Hamilton') female-biased sex ratio increase in frequency in the global population more than do females producing a non-biased ('Fisher') sex ratio. Second, in mixed groups containing both 'Fisher' and 'Hamilton' females, the 'Fisher' females produce more grandchildren than the 'Hamilton' females, as illustrated in the table. Since the population structure is as shown in the figure, this is a clear example of intrademic group selection.

An important feature of the table is that although relative fitness (as measured by the number of grandchildren produced) in mixed groups is higher for a foundress playing the 'Fisher' rather than the 'Hamilton' strategy, it nonetheless pays to play 'Hamilton' rather than 'Fisher' against either strategy. Whether a foundress ends up in a group with a 'Fisher' or a 'Hamilton', she will produce more grandchildren if she plays 'Hamilton' than if she plays 'Fisher'. 'Hamilton' is, after all, the evolutionarily stable sex ratio. Nevertheless, the group productivity approach to the problem reveals that foundresses pro-

ducing more daughters not only generate more mating opportunities for their sons (by both of the effects described by Taylor and Grafen) but also generate more mating opportunities for the sons of other females in the group. By producing a biased sex ratio, a 'Hamilton' female increases her own absolute fitness less than those of other group members who do not bias their sex ratios. This is not altruism because the 'Hamilton' female does not suffer any costs in terms of the number of grandchildren produced. In fact she produces

more grandchildren than if she herself had played the 'Fisher' strategy (see table).

The intrademic group selection approach adds a different perspective to previous discussions. For example, in Hamilton's model, it seems that the differential productivity of groups is always involved in the evolution of female-biased sex ratios^{9,16,17}. However, the female-biased sex ratio can evolve in the absence of discrete groups, if there is a tendency for local mate competition to occur in a global population with females distributed over a spatial continuum¹⁸. The need to distinguish actions performed for the good of the group from those that favour the individual is less relevant here than the need to realize that the spread of a gene depends on its fitness 'relative to others in the breeding population, and not to others with which it happens to interact'¹⁹. □

1. Maynard Smith, J. *Proc. R. Soc. B* 219, 315 (1983).
2. Hamilton, W.D. *Science* 156, 477 (1967).
3. Hamilton, W.D. in *Sexual Selection and Reproductive Competition in Insects* (eds Blum, M.S. & Blum, N.A.) 167 (Academic, New York, 1979).
4. Taylor, P.D. *Nature* 291, 64 (1981).
5. Grafen, A. in *Behavioural Ecology: an Evolutionary Approach* (eds Krebs, J.R. & Davies, N.B.) 62 (Blackwell, Oxford, 1984).
6. Charnov, E.L. *The Theory of Sex Allocation* (Princeton University Press, 1982).
7. Werren, J. *Evolution* 37, 116 (1983).
8. Maynard Smith, J. in *Behavioural Ecology* (eds Smith, R.H. & Sibly, R.) (British Ecological Society, 1984).
9. Colwell, R.K. *Nature* 290, 401 (1981).
10. Wilson, D.S. *A. Rev. ecol. Syst.* 14, 159 (1983).
11. Wynne-Edwards, V.C. *Animal Dispersion in Relation to Social Behaviour* (Oliver & Boyd, Edinburgh, 1962).
12. Wright, S. *Ecology*, 26, 415 (1945).
13. Maynard Smith, J. *Nature* 201, 1145 (1964).
14. Maynard Smith, J. *Q. Rev. Biol.* 51, 277 (1976).
15. Wade, M.J. *Q. Rev. Biol.* 53, 101 (1978).
16. Colwell, R.K. *Nature* 298, 495 (1982).
17. Wilson, D.S. & Colwell, R.K. *Evolution* 35, 882 (1981).
18. Charlesworth, B. & Toro, M.A. *Nature* 298, 494 (1982).
19. Maynard Smith, J. *Br. ecol. Symp.* 23, 195 (1984).

Paul H. Harvey is in the Biology Department, Princeton University, New Jersey 08540; Linda Partridge is in the Department of Zoology, University of Edinburgh, Edinburgh EH9 3JT, UK; and Len Nunney is in the Biology Department, University of California, Riverside, California 92521, USA.

Archaeology

Scandinavian pine chronologies

from John Fletcher

LYING close to entrances from the Baltic to the Gulfs of Bothnia and Finland, the Åland islands have played the role in the Baltic that Malta has played in the Mediterranean. The massive castle of Kastelholm on the main island was already of strategic importance in the 14th century. Now its importance is more to archaeologists, including those at a recent conference* held in Mariehamn, the main town of the autonomous province of Åland.

The granite promontory on which the castle stands was originally surrounded by a palisade of pine stakes. Most of the stakes that still exist are medieval and so of great interest to dendrochronologists in a region that is beyond the habitat of oak. To form pine chronologies for Åland, L. Löpstrand (Tree-Ring Laboratory, Uppsala) has used not only the palisade stakes but also boards and planks from the castle itself and churches on the islands. Elsewhere in the north and east of Scandinavia, mediaeval excavations at Trondheim, Bergen and Oslo, in Norway, have provided T. Thun (Botanical Institute, Trondheim) with pine samples, while a mean curve for central Sweden formed by T. Bartholin (Institute for Wood Anatomy, Lund) goes back to AD 1078; Bartholin's samples for Lapland go back as far as AD 436.

So far, Scandinavian pine-ring dendrochronologists have not attempted to use any long-distance (over 500 km) correlations such as those applied to sub-alpine conifers and to oak in central Germany and northwest Europe. The almost obsolete concept that each chronology needs to be regional and confined to small areas — it was thought even to the Åland islands themselves — persists in Scandinavia and has delayed progress in dating by dendrochronology. A link between the growth of pine in west Norway and that measured by Schweingruber, Bräker and Schär (*Boreas* 8, 427; 1979) in northern Scotland is no more unlikely than the link that has been established between the growth of oak in certain periods in southern England and Denmark. For Åland and southern Finland, the medieval pine chronology from AD 884–1462 formed by B.A. Kolchin (*Soviet Archaeology* 1, 113; 1962)[†] from multilayers of the mediaeval wooden pavements of Novgorod (2,500 samples in 1961) may prove useful because trees from forests many hundreds of kilometres apart show the same 'thin-ring' characteristics.

Also of considerable interest at the meeting was some information on the diet of prehistoric man derived from stable isotope analysis of bones of Danish skeletons (H. Tauber, National Museum, Copenhagen). For northern latitudes, the decrease in the ratio of ^{13}C to ^{12}C when carbon dioxide enters the ocean is substantially less than when it is assimilated into plants via photosynthesis, so a fish-based diet can be distinguished from a plant-based diet or

one that is based on herbivorous animals. Human bones, and those of dogs, examined by Tauber show a local change in subsistence in the long mesolithic period which covers the Maglemose and Ertebölle stone age cultures (~7000–3000 BC, uncalibrated) in Scandinavia. In the earlier Maglemose culture, almost no fish was consumed but subsequently fish played a large part in the diet (so Tauber refers to the 'hunter' and 'fisher' Stone Age periods). Nitrogen isotope ratios ($^{15}\text{N}/^{14}\text{N}$) confirm the results from carbon isotope analysis. □

John Fletcher is in the Research Laboratory for Archaeology and the History of Art, Oxford OX1 3QJ, UK.

Palaeontology

Ammonoids and extinctions

from Digby J. McLaren

SPECULATION on causes of extinctions in the geological past dates back to the early days of geology when Cuvier's catastrophic model was replaced by Lyell's gradualism, giving rise to Whewell's derisory term 'uniformitarianism'. Today several immediate causes are fielded, both gradual and sudden, and there is a considerable degree of polarization among their protagonists. Because of the nature of the fossil record, most hypotheses are concerned principally with marine faunal extinctions, although they may be generalized to the land and the plant kingdom. They include climatic change¹, anoxic oceanic events², eustatic events, either a fall³ or a rise^{4,5} in sea level, and large-body impacts^{6,7}. It seems increasingly likely that a variety of mechanisms, acting singly or severally, have caused environmental change leading to extinctions at different times in Earth history. Furthermore, confusion between conflicting hypotheses arises through failure to distinguish between immediate or proximate causes and the three possible ultimate causes: internal terrestrial effects (for example, mantle movements), orbital and solar changes, and random or episodic meteoroid or cometary impacts⁸.

On page 17 of this issue, M.R. House has used his knowledge of the Ammonoidea, an extinct subclass of the cephalopods, in an attempt to use evolutionary changes as a 'bioseismograph' to detect events causing environmental change⁹. He considers eight extinctions over a selected period of ~30 Myr from the Lower Devonian to the base of the Carboniferous. He recognizes a sequential pattern in each of these events — a decline in diversity and an inferred reduction in biomass, followed by appearance of new forms, increase in diversity and the re-establishment of normal evolutionary trends. Interestingly, similar patterns have been reported for extinctions of other groups including trilobites in the Cam-

brian¹⁰, mammal-like reptiles from Carboniferous to Trias¹¹, and Foraminifera at the beginning of the Paleocene¹², as well as the abrupt extinction and slow subsequent re-establishment and radiation of most corals and brachiopods after the end of the Frasnian^{13,14} (which House calls the Kellwasser event).

The eight extinction events differ in degree. Some are now recognizable only by detailed changes in ammonoid development; three are large and known to affect many unrelated groups:

(1) The Taghanic event, which marked the end of the Givetian (according to established general usage — House rightly deplores the changes made by the Devonian Subcommittee¹⁵), was a major extinction, particularly well defined by marked changes in the ammonoids. It represents the beginning of the end of the eastern North American faunal province.

(2) The Kellwasser event (the name of which is, in my opinion, unfortunately chosen for a world-wide extinction at the end of the Frasnian). Space does not allow development of the argument, but Walliser⁵ and House¹⁶ himself have shown that euxinic black shale deposition, which marks the beginning of this event in Germany, began before the end of the Frasnian. Typical Frasnian corals and brachiopods flourished in other parts of the world, for example Alberta and north-west Australia, and became extinct at about the end of the *Palmatolepis gigas* conodont zone. The Famennian fauna of the early Stage was depauperate and markedly different, and the Stage saw the slow development of new radiations in several groups that were the precursors of the Carboniferous fauna. Recently an iridium anomaly has been discovered in the Upper Devonian of the Canning Basin, north-west Australia, at the horizon of the late Frasnian extinction event¹⁷. Although the

*The 3rd Nordic Conference on 'The Application of Scientific Methods to Archaeology', Mariehamn, 8–11 October, 1984, was organized by Dr Högne Jungner of the Radiocarbon Dating Laboratory, Helsinki. The Proceedings will form a special issue of ISKOS, published by the Archaeological Society of Finland.

†English translation in Fletcher J.M. & Linnard W. *Russian Papers on Dendrochronology and Dendroclimatology 1962, 1968, 1970, 1972*, 124 (Oxford University, 1977).

trapping mechanism is unusual, nevertheless the anomaly suggests the possibility that a meteoroid impact or major volcanic event has been a contributing cause to the world-wide extinction¹⁸. There are still problems to be settled, however, particularly in regard to the synchronicity of the event in terms of conodont zonation.

(3) The Hangenberg event, towards the close of the Famennian, was another euxinic black shale event in Europe that caused a relatively extended decline and extinctions which included the end of the peculiar Devonian ammonoid group, the clymeniids. The coral and brachiopod faunas of the late Famennian are nevertheless more closely related to succeeding Carboniferous forms than to any previously existing in the Frasnian.

While House⁹ suggests that the events he describes may have a eustatic or euxinic immediate cause, he is prepared to speculate on the possibility of various ultimate causes. Recently he has published on Devonian eustatic events, and has attempted to recognize rhythms and cycles in the New York Devonian and their correlates across the Atlantic, based on sedimentological and biostratigraphical data¹⁶. In the current paper, he finds that facies micro-rhythms may, in some instances, be correlated with Milankovitch cycles of the order of 100 Kyr. Other possibilities, including climatic change and carbon variations, would seem to make a simple rhythmic interpretation for all the observed events unlikely. Nor is it easy to decide if larger-scale periods are involved. At least some of the events described by House do not seem to be regularly spaced and there is variation in the magnitude between them. He notes recent ideas linking periodicity in extinctions with rhythmic episodes of enhanced meteoroid or cometary impact, but he concludes that, at least in the Devonian and Carboniferous, rhythms and events did not exhibit periodicity. His evidence suggests a more complex scenario and he concludes that thorough investigation of the type of events he describes might lead to an accurate chronology and ultimately to an objective global time scale. □

1. Stanley, S.M. *Geology* **12**, 205 (1984).
2. Fischer, A.G. & Arthur, M.A. *Soc. econ. Paleont. Mineral spec. Publ.* **25**, 19 (1977).
3. Hallam, A. *Rev. Earth Planet. Sci.* **12**, 205 (1984).
4. House, M.R. *Proc. Yorks. geol. Soc.* **40**, 459 (1975).
5. Walliser, O.H. *Terra Cognita* **4**, 17 (1984).
6. McLaren, D.J. *J. Paleont.* **44**, 801 (1970).
7. Alvarez, L.W., Alvarez, W., Asaro, F., & Michel, H.V. *Science* **208**, 1095 (1980).
8. McLaren, D.J. *Geol. Soc. Am.* **94**, 313 (1983).
9. House, M.R. *Nature* **313**, 17 (1985).
10. Palmer, A.R. *J. Paleont.* **58**, 599 (1984).
11. Kemp, T.S. *Mammal-like Reptiles and the Origins of Mammals* (Academic, New York, 1982).
12. Smit, J. *Geol. Soc. Am. spec. Pap.* **190**, 329 (1982).
13. Pedder, A.E.H. *Geol. Soc. Am. spec. Pap.* **190**, 485 (1982).
14. Boucot, A.J. *Evolution and Extinction Rate Controls* (Elsevier, Amsterdam, 1975).
15. Ziegler, W. & Klapper, G. *Episodes* **4**, 18 (1984).
16. House, M.R. *Proc. Usher Soc.* **5**, 396 (1983).
17. Playford, P.E., McLaren, D.J., Orth, C.J., Gilmore, J.S. & Goodfellow, W.D. *Science* **226**, 437 (1984).
18. McLaren, D.J. *Geology* (in the press).

Digby J. McLaren is in the Department of Geology, University of Ottawa, Ontario K1N 6N5, Canada.

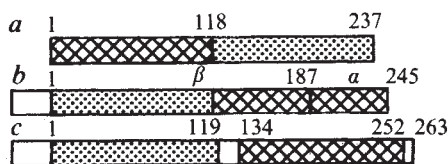
Plant biochemistry

Seed protein construction

from John A. Gatehouse and Donald Boulter

PRODUCTION of seed proteins by the enzymatic cleavage of precursor molecules into their component parts is common place. But on page 64 of this issue, D.M. Carrington, A. Auffret and D.E. Hanke make the provocative suggestion that concanavalin A (con A), a seed protein of jack-bean (*Canavalia ensiformis*), is produced by the rejoining, in transposed order, of the two main cleavage products of its precursor¹.

Con A belongs to the class of carbohydrate-binding proteins known as lectins. These are specifically accumulated in the seeds of many plants, although usually to a lesser extent than the storage proteins, and may function in the defence of seeds against insects and other predators. Lectins have been purified from a range of legume seeds and a considerable body of data has shown that the lectins of pea, broad bean, lentil and related species have highly related amino acid sequences. The sequence of con A is also related if a circular permutation of sections of its sequence is taken into account^{2,3}. Specifically, pea lectin is synthesized as a precursor containing a leader sequence, followed by its 187-amino acid β subunit and then its 58-amino acid α subunit⁴, whilst con A is a tetramer of a 237-amino acid polypeptide, of which amino acids 1-118 are homologous to 105-240 of pea lectin, and amino acids 119-237 are homologous to 1-105 of pea lectin (see figure).



Schematic structure of con A (a), pea lectin (b) and the suggested con A precursor (c). Corresponding regions are in the same shade. Numbers refer to amino acids.

Enzymatic cleavage of con A between amino acids 118 and 119 (ref.5), and possibly at other positions too⁶, has seemed to be a slow, but natural, modification, accounting for the frequent isolation of the two fragments of the lectin. But Carrington *et al.* now present evidence that the cleaved fragments are, in fact, precursors to a 237-amino acid protein that is formed by polypeptide ligation. (The isolation of fragments is thus ascribed to incomplete ligation). This interpretation is based on the isolation of a jack bean seed cDNA (produced from mRNA) that seems to contain the complete con A coding sequence but rearranged in a way that makes it similar to the sequence of pea lectin (see figure). Specifically, the cDNA encodes a 29-amino acid leader sequence, followed

by amino acids 119-237 of con A, a 15-amino acid 'linker' peptide, amino acids 1-118 of con A, and a 9-amino acid carboxy-terminal peptide. To produce concanavalin A from this precursor would require at least four proteolytic cleavages and the ligation of the 1-118 fragment to the 119-237 fragment. At least one further proteolytic step must be involved, since a form of con A isolated from immature seeds contains an extra four amino acids corresponding to the last few of the linker fragment at its amino-terminus. In addition, the linker fragment carries a glycosylation site, so that the precursor is probably glycosylated whereas con A itself is not.

Most of this sequence of events is not unreasonable when compared to the synthesis of other seed proteins, particularly since, as with other legume seed proteins⁷, all the proteolytic cleavages (except that of the leader sequence) occur on the carboxy-terminal side of asparagine. But the postulated final ligation step is unprecedented and should be treated with caution. Is it possible that Carrington *et al.* have sequenced a cDNA corresponding not to con A but to a related protein or to a precursor that gives rise to the 'fragmented' subunits seen *in vivo*, whereas another cDNA encodes the 'intact' subunits? The authors think not, since they claim that only one con A gene per genome can be detected with their cDNA, although complete evidence is not presented. They also claim that the precursor form of con A predicted by the cDNA sequence is detected *in vivo* as a polypeptide antigenically-related to con A, and can be shown to give rise to all the con A polypeptides seen *in vivo*. Again, however, they provide only incomplete evidence (pulse-chase data) that this is the case. Furthermore, the possible presence of other sites of proteolytic cleavage is not considered.

Finally, one must ask why such a per-verse synthetic route should be used. And if it is used, what are the characteristics of the ligation enzyme? One suspects this system still has a few tricks up its sleeve. □

1. Carrington, D.M., Auffret, A. & Hanke, D.E. *Nature* **313**, 64 (1985).
2. Foriers, A., Lebrun, E., van Rapenbusch, R., De Neve, R. & Strosberg, A.D. *J. Biol. Chem.* **256**, 5550 (1981).
3. Hemperly, J.J. & Cunningham, B.A. *Trends biochem. Sci.* **8**, 100 (1983).
4. Higgins, T.J.V., Chandler, P.M., Zwawski, G., Button, S.C. & Spencer, D. *J. Biol. Chem.* **258**, 9544 (1983).
5. Alec, Y., Iwakuchi, M. & Ishii, S.I. *Biochem. biophys. Res. Commun.* **45**, 1271 (1971).
6. Dalkin, K., Marcus, S. & Bowler, D.J. *Planta* **157**, 531 (1983).
7. Gatehouse, J.A., Croy, R.R.D. & Boulter, D. *CRC Crit. Rev. Pl. Sci.* **1**, 287 (1984).

John A. Gatehouse and Donald Boulter are in the Department of Botany, Durham University, Durham DH1 3LE, UK.

Treaty verification

Seismic noise and a test ban

from Thomas C. Bache

ONE of the major obstacles to agreement on a comprehensive test ban treaty (CTBT) on nuclear explosions is the concern about 'verification' — that is, the ability to detect and identify clandestine explosions. There has been a continuing debate among Western seismologists about the possibility of verification through a network of seismic stations located entirely outside the Soviet Union, but the consensus is that a significant clandestine testing programme could be concealed from such a network. Therefore, the United States and United Kingdom have always insisted that stations inside the Soviet Union are required for a verifiable treaty. For many years the Soviet Union rejected this position, but in the 1977–1980 negotiations, it agreed, in principle, to allow internal seismic stations for verification purposes, although many contentious issues about the numbers, locations and design of these stations remained unresolved when the negotiations went into recess in November 1980.

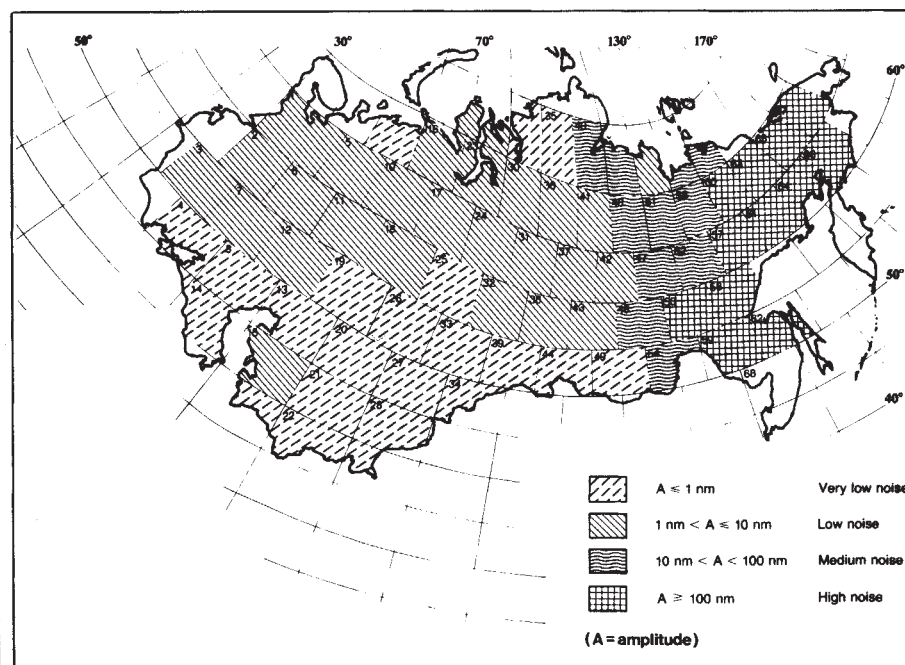
The most important advantage of internal stations is that they increase the network sensitivity; that is, they lower the threshold for detection of small events (or small signals from disguised explosions). Seismologists have developed simulation programs that map the detection threshold for different seismic network configurations^{1,2}, and the output of these programs depends on assumptions about signal amplitude decay as a function of distance and about the ambient 'noise' level at the stations (detection is essentially a function of the signal-to-noise ratio). Seismic noise is the ground motion background, out of which must be picked the discrete signals from earthquakes and explosions. Because it varies from region to region and site to site, actual data are needed to estimate the noise background even to within an order of magnitude, but there is an immense lack of reliable data from within the Soviet Union. In a report³ published in 1984, P.W. Rodgers and A.J. Piwinski of Lawrence Livermore National Laboratory have compiled the best available noise information from the Soviet literature.

The report includes an appendix with brief summaries of the twenty publications found to contain relevant information. Unfortunately, only a few of these give any useful data, and the data are very difficult to interpret because there is little description of how the noise measurements were made. (Attempts to resolve the problems by correspondence met with no success.) The most useful data on noise were found in a monograph by Rykunov⁴ which gives noise estimates at 4, 5, 6 and 7 second periods for 3×10^5 km² quadrangles

covering the entire country. Since the monograph does not say how these numbers were determined, Rodgers and Piwinski assume they are spatially and temporally averaged peak-to-peak amplitudes in micrometers. Furthermore, the character of the tabulated values suggests that they are based on some interpolation or smoothing of values determined from seismograms recorded on paper. The most useful data from the internal stations are likely to be at frequencies of 1 Hz and more, so the most relevant results in the report appear on the 1-second noise map (reproduced here). The map was obtained by extrapolating from the Rykunov values at 4, 5 and 6 seconds, which is not an

America⁵. Moreover, stations in adjacent countries (including Iran, Afghanistan and Turkey) are not so quiet.

The difficulty of estimating noise levels is typical of those encountered in addressing nearly all of the geophysical issues involved in estimating monitoring capabilities or in deciding whether or not Soviet explosions have exceeded the 150 Kt limit of the Threshold Test Ban Treaty. There is a vast open literature on nuclear explosions and earthquakes occurring in the United States and recorded at internal and external stations. Digital seismograms of such events and of station noise are openly available from the US Geological Survey. In contrast, though the Soviet Union apparently has an extensive research programme in seismology, few results and almost no usable data appear in open sources. As an example, while many Soviet stations routinely report the arrival time and amplitude of detected signals to the International Seismic Center in England, data on signals



unreasonable procedure given the paucity of the information, but is certainly controversial.

The map seems to represent the best noise data available from within the Soviet Union but, unfortunately, will be of little help in making confident estimates of the ability to verify a CTBT. One reason is that the map gives so little information: very large regions are represented by a single value and we can only guess what this value actually means. Beyond that there are some surprising features of the map that are difficult to believe without corroborating evidence. For example, the range of 'average' noise spans more than two orders of magnitude, much more than on apparently comparable maps of North America. Also, for huge areas, the average noise at one-second periods is less than 1 nanometer. Such a low level of noise has only been observed at specially selected sites in North

from nuclear explosions are always deleted from the reports. Thus, as far as a CTBT is concerned, poorly supported assumptions based on data recorded outside the Soviet Union are required to estimate the potential capability of stations within it. This is a major reason for the large differences in published estimates^{2,5} of how effective such stations would be in assuring Western nations that the Soviet Union was actually complying with a comprehensive test ban treaty. □

1. Evernden, J.F. *Bull. Seis. Soc. Am.* **66**, 281 (1976).
2. Hannon, W.J. *Ener. & Tech. Rev. UCRL - 5200 - 83-5*, 50 (1983).
3. Rodgers, P.W. & Piwinski, A.J. *Seismic Noise in the U.S.S.R. UCRL - 53328* (Lawrence Livermore National Laboratory, 1982).
4. Rykunov, L.N. *Microseisms, Experimental Characteristics at Periods 0.07-8 Seconds* (Nauka Press, Moscow, 1967).
5. Sykes, L. & Evernden, J.F. *Scient. Am.* **247**, 47 (1982).

Thomas C. Bache is at the Science Applications International Corporation, 10210 Campus Point Road, San Diego, California 92121, USA.

Earth sciences

Statistical structure of geomagnetic reversals

from D. V. Kent

REVERSALS of the Earth's magnetic field have not occurred at a constant average rate over the past 170 Myr^{1,2}. Shifts in the mean reversal rate at various times appear to be separated by periods of tens of millions of years in which the average rate of reversals remains nearly constant. Conclusions drawn from these and other statistical properties of reversal processes have been based on analyses of early versions of the geomagnetic polarity timescale (GPTS). More recent versions of the GPTS for the past 170 Myr incorporate a variety of new data from marine magnetic anomalies, magneto-biostratigraphic correlations and improved age calibrations, resulting in a new picture of the reversal behaviour of geomagnetic fields^{3,4}.

McFadden and Merrill³ have analysed the mean rate of reversals using a sliding window with a fixed number of reversals. This reveals an approximately linear trend, increasing from less than one reversal per million years at about 86 Myr BP (just after the 30 Myr interval of constant normal polarity in the Cretaceous) to over four per million years for the past few million years. Lowrie first arrived at a similar trend using a sliding window of fixed duration⁴. Then he and I documented such reversal behaviour in a different revised GPTS⁵. McFadden and Merrill suggest a change in the structure of the rate variations at about 10 Myr BP, although it has been suggested^{4,6} that several maxima in reversal frequency associated with an apparent 15 to 20 Myr fluctuation are superposed on the long term linear trend.

The linear trend for the past 86 Myr is suggested by McFadden and Merrill to be part of a broader structure of the variable reversal process. The overall pattern seems to show a regular decline in reversal frequency from 165–119 Myr BP, when the process evidently ceased and the field remained in the same polarity state for ~30 Myr; then from about 86 Myr BP onward, field reversals again occurred at a gradually increasing rate. The observed long-term changes in mean reversal rate may represent secular changes in conditions in the Earth's core. McFadden and Merrill believe that variations in the mean temperature at the base of the mantle over long time periods are the most likely cause of the changes. Their interpretation requires at least intermittent convection deep within the mantle and it suggests that temperatures at the core-mantle boundary are highest during the long periods of constant polarity.

Despite the various refinements made to the GPTS, a number of important prob-

lems remain to be resolved before much confidence can be placed in our knowledge of the statistical structure of reversals. One long-standing but central issue concerns the presence or absence of short polarity intervals. It is clear that the distribution of observed polarity intervals is generally too deficient in short intervals for it to be fully compatible with either an exponential distribution or a Poisson reversal process in which the probability for the next reversal to occur is constant with time. Instead, it has been shown^{2,7} that a gamma distribution provides a good fit to the observed polarity intervals, at least for the past 86 Myr.

In a companion paper to reference 3, McFadden argues that the observed distribution embodies a Poisson process combined with a filtering process that preferentially removes the record of short polarity intervals, giving rise to the observed gamma distribution⁷. Accordingly, McFadden and Merrill³ calculate that the actual reversal frequency was 46 per cent higher than is observed, so that about half of the observed polarity intervals represent a concatenation of more than one interval of the actual reversal sequence; magnetostratigraphers, take note. The alternative is that short intervals seldomly occurred. This would imply a renewal process in which the probability per unit time of a reversal is not constant and that the occurrence of each reversal is not independent of earlier ones.

Difficult though it is to obtain firm evidence of short polarity intervals, because they are at or below the threshold of resolution of most palaeomagnetic studies, it will be important to do so because of their dramatic effect on the statistics of geomagnetic reversals according to our calcul-

ations^{4,8}. Small-scale marine magnetic anomalies ('tiny wiggles') have been observed over high spreading-rate ridge systems and among other possibilities, may represent short (<40 kyr) polarity intervals; because of their uncertain nature, 'tiny wiggles' are usually not uniformly included in the GPTS⁸. If, however, tiny wiggles are all taken to represent ~30 kyr polarity intervals, they have a dramatic effect on the polarity-interval distribution, particularly over the past 40 Myr, in which 50 tiny wiggles have been identified. Taking them into account, the overall reversal frequency is increased by almost 80 per cent, and the sliding mean reversal rate becomes constant over this time span⁴. Although the distribution of all polarity interval lengths then approximates to an exponential, the normal and reverse polarity interval distributions develop an asymmetry in statistical properties, due to the unequal distribution of the added short intervals which are overwhelmingly of normal polarity.

If little less, this exercise with problematic tiny wiggles serves to re-emphasize how strongly the existence of short polarity intervals influences the statistical structure of geomagnetic polarity reversals. There is also the serious problem that spurious increases in reversal frequency may correspond to the most intensively studied time intervals. In my opinion a 'more' reliable description of the reversal process, whether it actually shows polarity asymmetry or even long-term linear trends in reversal frequency, must continue to await further improvements in resolution of the geomagnetic reversal timescale. □

1. Cox, A. *Rev. geophys. space Phys.* **13**, 35 (1975).

2. Phillips, J.D. *J. geophys. Res.* **82**, 835 (1977).

3. McFadden, P. & Merrill, R. *J. geophys. Res.* **89**, 3354 (1984).

4. Lowrie, W. & Kent, D. *Earth planet. Sci. Lett.* **62**, 305 (1983).

5. Lowrie, W. *Phil. Trans. R. Soc. A* **306**, 129 (1983).

6. Mazaud, A., Laj, C., de Seze, L., & Verosub, K.L. *Nature* **304**, 328 (1983).

7. McFadden, P.L. *J. geophys. Res.* **89**, 3363 (1984).

8. LaBrecque, J., Kent, D. & Cande, S. *Geology* **5**, 330 (1977).

D. V. Kent is in the Lamont-Doherty Geological Observatory and Department of Geological Sciences, Columbia University, Palisades, New York 10964, USA.

Molecular biology

Sigma factors in multitude

from Andrew Travers

SIXTEEN years ago, almost to the day, the discovery of a sigma initiation factor for the bacterium *Escherichia coli* was reported in *Nature*¹. The factor is a polypeptide that binds to the enzyme RNA polymerase, conferring upon it the ability to initiate the transcription of certain genes. A conglomeration of recent papers illustrates the striking ramifications of that initial report, making it clear that sigma factors are widespread among bacteria, and that they are related to the polypeptide which mediates

the response of bacteria to a sudden increase in temperature (heat shock). Moreover, it becomes possible to suggest how the different mechanisms that mediate the analogous responses to heat shock in bacterial and eukaryotic cells have evolved.

In *E. coli*, sigma factor binds to the 'core' RNA polymerase to form a holoenzyme that initiates transcription accurately from promoters that conform to the '-35' and '-10' consensus sequences of most *E. coli* promoters. It has long been

hypothesized that bacteria contain multiple sigma factors, each of which directs the core RNA polymerase to initiate at a different class of promoters but *Bacillus subtilis* has remained the only organism for which this principle of gene regulation has been demonstrated. Now, however, on page 22 of this issue Westpheling *et al.* report the identification of multiple holoenzyme forms in the related, but more complex, bacterium *Streptomyces coelicolor*⁵. One form, containing a sigma factor of relative molecular mass 35,000 (35K), has the same promoter specificity as the holoenzymes of vegetatively-growing *E. coli* and *B. subtilis*, an observation that neatly complements the discovery of such a promoter site in *S. lividans* (M. Buttner and N. Brown, personal communication). The second form of *S. coelicolor* holoenzyme contains polypeptides of 49K and 37K, each of which independently stimulates transcription from a promoter utilized by the *B. subtilis* holoenzyme with a homologous 37K factor.

The probable general occurrence of multiple sigma factors in bacteria is apparent from recent elegant studies of the heat-shock response of *E. coli*, in which a sudden increase in temperature stops the production of vegetative gene transcripts⁶ and starts the transcription of 15 heat-shock genes which require the product of the *htpR* (*hin*)^{7,8} gene for their expression. Landick *et al.*⁹ have recently sequenced this gene and shown that it has strong homology in two regions with the sigma factor of vegetative *E. coli*. In agreement with the original sigma hypothesis, one of these domains has the characteristics of a DNA-binding site while — the authors suggest —

the second is involved in making contacts with the core RNA polymerase. In a complementary study Grossman *et al.*¹⁰ have demonstrated that the *htpR* gene product functions as a sigma factor *in vitro*, binding to the core RNA polymerase and directing initiation at a heat-shock promoter not utilized by the vegetative holoenzyme.

The identification of multiple sigma factors in bacteria has proceeded by two different routes. In *B. subtilis* and *Streptomyces* the biochemical characterization of factors of unknown biological function came first and is only now being supplemented by genetic studies such as those of Stragier *et al.*¹¹ who have sequenced the *spoIIG* gene, a locus required for sporulation, and shown that it has strong homology to the sigma factor of vegetative *E. coli* in regions corresponding to those observed for the *htpR* gene. By contrast, the identification of regulatory loci such as *htpR* of *E. coli* and gene 55 of the T4 bacteriophage¹² preceded their functional assignment as sigma factors.

Are there sigma factors still to find? The answer is probably yes. For example, the nitrogen fixation (*nif*) genes in *Klebsiella* and *Rhizobium* contain a highly conserved promoter sequence which is completely unrelated to those of the other known promoters^{13,14}, suggesting that either or both of the products of the regulatory *nifA* and *glnF* genes function as sigma factors.

How does one sigma factor functionally replace another? Activation of heat-shock genes by the gene product depends critically on the *in vivo* concentration of the vegetative sigma factor¹⁵, suggesting that the two factors compete for the core polymerase. In this situation, any mechanism that alters the relative affinity of the factors for the core polymerase might be sufficient to effect the switch. Another trigger for the synthesis of heat-shock proteins in *E. coli* is amino acid starvation¹⁰, which also results in the inactivation of many promoters transcribed by the vegetative holoenzyme. The molecule that mediates this control is the nucleotide 'alarmone' guanosine 5'-diphosphate 3'-diphosphate (ppGpp), which directly affects polymerase function. Heat shock induces the accumulation of another alarmone, AppppA¹⁶, synthesized in an aberrant reaction by aminoacyl tRNA synthetases¹⁷. Perhaps such alarmones act by loosening the protein-protein interactions between the vegetative sigma factor and core RNA polymerase, thereby concomitantly switching off a major class of vegetative genes and facilitating the replacement of one sigma factor by another.

The major 70K heat-shock protein is one of the few proteins known to be highly conserved between eukaryotes and eubacteria¹⁸, and the suggested antiquity of the heat-shock response makes it an obvious place to look for clues to the evolution of transcriptional systems. Consequently, it is striking that the '-35' consensus sequence (CNNCTTGAA) of the heat-shock res-

ponsive promoters of *E. coli* (but not necessarily all bacterial) genes contains the first half of the inverted repeat consensus sequence (CTNGAANNNTTCNAG) of eukaryotic heat-shock responsive genes¹⁹. Recognition of the latter is by a DNA-binding protein^{21,22} and of the former by a sigma factor. How could such apparently disparate mechanisms have evolved? A simple hypothesis would be that the primeval RNA polymerase, like the present day core of RNA polymerase II, possessed a predilection for melted DNA but lacked sequence-specific recognition, which was conferred by a protein that bound both to the polymerase and to a DNA sequence in the vicinity of easily melted DNA sites. In the evolution of eubacteria, the affinity of this specifying protein for polymerase became greater than that for DNA. By contrast, in the course of eukaryotic evolution, the specifying protein came to have a greater affinity for DNA than polymerase though retaining the ability to interact with it²³. These considerations imply that the '-35' region recognized by sigma factors must remain asymmetric to position the enzyme in the correct orientation to melt the DNA in the '-10' region. In the case of eukaryotic polymerase II promoters, the constraint for directionality would be removed by the one or more specific proteins that identify the corresponding TATA box. This would allow more flexibility in the structure of the specifying region and thus permit multiple binding sites, directing the cooperative binding of more than one molecule of the corresponding regulatory protein (for example, in *Drosophila* hsp 70 promoters²⁴) or of distinct regulatory proteins with different sequence specificities (for example, in *Xenopus* hsp 70 promoters²⁵). □

100 years ago

EARTHQUAKES IN ENGLAND, AND THEIR STUDY

As no record of the most recent earthquake shock in England has yet found a place in the pages of *Nature*, perhaps I may be permitted to give the following slight details, collected from the daily papers of Lancashire and London for November 15:—

A shock of earthquake ("severe," yet causing no structural injury) was experienced at Clitheroe, and in the neighbourhood, on the evening of November 14. At about 5.10 p.m. a terrific report, resembling loud thunder, was heard, instantly followed by a strong vibration of the earth, sufficient to induce the inhabitants to run out of their houses into the streets in a terrified state. Much excitement prevailed throughout the borough and neighbourhood of Clitheroe, especially at Low Moor.

The circumstance that this particular part of Lancashire is much subject to earthquake disturbances, makes it specially important that no details of their occurrence be lost to science. Within the last fifty years at least six well-authenticated shocks have been recorded,—in 1835, 1843, 1868, 1871, 1873, and 1884,—and this list might easily be extended. Lancashire, indeed, may be considered as one of the chief areas of disturbance in England, and after Comrie, in Perthshire, perhaps the most important centre of seismic action in Great Britain. From *Nature* 31, 172, 25 December 1884.

1. Burgess R.R. *et al. Nature* 221, 43 (1969).
2. Johnson, W.C., Moran, C.P. Jr & Losick, R. *Nature* 302, 800 (1983).
3. Doi, R.H. *Archs Biochem. Biophys.* 214, 772 (1982).
4. Gilman, M. & Chamberlin, M. *Cell* 35, 284 (1983).
5. Westpheling, J., Ranes, M. & Losick, R. *Nature* 313, 22 (1985).
6. Ryals, J., Little, R. & Bremer, H. J. *Bacteriol.* 151, 1425 (1982).
7. Neidhardt, F.C. & Van Bogelen, R.A. *Biochem. biophys. Res. Commun.* 100, 894 (1981).
8. Yamamoto, T. & Yura, T. *Proc. natn. Acad. Sci. U.S.A.* 79, 860 (1982).
9. Landick, R. *et al. Cell* 38, 175 (1984).
10. Grossman, A.D., Erickson, J.W. & Gross, C.A. *Cell* 38, 383 (1984).
11. Stragier, P. *et al. Nature* 312, 376 (1984).
12. Kassavetis, G.A. & Geiduschek, E.P. *Proc. natn. Acad. Sci. U.S.A.* 81, 5101 (1984).
13. Benyon, J. *et al. Cell* 34, 665 (1983).
14. Ausubel, F.M. *Cell* 37, 5 (1984).
15. Osawa, T. & Yara, T. *Mol. gen. Genet.* 184, 166 (1981).
16. Lee, P.C., Bochner, D.M. & Ames, B.N. *Proc. natn. Acad. Sci. U.S.A.* 80, 7496 (1983).
17. Rapaport, E., Srihovec, S.K. & Zamecnik, P.C. *Proc. natn. Acad. Sci. U.S.A.* 72, 2653 (1975).
18. Bardwell, J.C.A. & Craig, E.A. *Proc. natn. Acad. Sci. U.S.A.* 81, 848 (1984).
19. Pelham, H.R.B. *Cell* 30, 517 (1982).
20. Cowing *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
21. Parker, C.S. & Topol, J. *Cell* 37, 253 (1984).
22. Wu, C. *Nature* 309, 229 (1984).
23. Travers, A. *Nature News and Views* 303, 755 (1983).
24. Dudler, R. & Travers, A.A. *Cell* 38, 391 (1984).
25. Bienz, M. *Trends biochem. Sci.* (in the press).

Andrew Travers is at the Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

Correlation of mid-Palaeozoic ammonoid evolutionary events with global sedimentary perturbations

Michael R. House

Department of Geology, University of Hull, Hull HU6 7RX, UK

Eight ammonoid extinction events occurring within 30 million years seem to be correlated with sedimentary changes in global sea level and temperature. These evolutionary patterns may be regarded as a bioseismograph reflecting events causing environmental change.

THE Ammonoidea provide good examples of the discontinuous nature of evolution often observed in the fossil record. Coiled ammonoids survived for about 330 million years from the Lower Devonian to the end of the Cretaceous (ref. 1; Fig. 1). Major extinction events during ammonoid evolution have been used for time subdivisions and have been correlated with palaeogeographical change (suggested in refs 2,3).

Finer-scale evolutionary fluctuations gave rise to the refined time scale in current use; these fluctuations define stufen divisions in the Devonian, genus zones of the late Palaeozoic and 'ages' in the Mesozoic. These short periods were characterized by 'morphological packets' of distinctive forms closely followed by the next successful packet, resulting in an evolutionary-relay effect^{3,4} which helps to define the time unit. Changes of yet lower taxonomic order define zonal and subzonal boundaries. Correlation of these finer-scale evolutionary events with environmental and palaeogeographic change is less well documented⁵⁻⁷.

I consider here eight examples of extinction events in the mid-Palaeozoic history of the Ammonoidea from the Lower Devonian to basal Carboniferous, a time span of ~30 Myr. Each example is usually characterized by six features: (1) some decline in diversity in the earlier morphological packet; (2) a short period of reduced diversity and inferred biomass; (3) the cryptic appearance of novel groups diagnosed by changes in early ontogeny; (4) an increase in taxonomic diversity and abundance of the new stocks; (5) a pattern in the new groups showing gradual evolutionary trends usually towards morphological complexity and increased size; and (6) a correlation between the initiation of decline (feature 1) or of the low point (feature 2) with euxinic conditions, often associated with global transgressive pulses.

The classic New York Devonian succession (Fig. 2) shows the international stage, conodont and ammonoid terminology and inferred local depth changes, many of which appear to correlate with international eustatic events reviewed elsewhere⁸. In the present study, for simplicity, names replace letters used previously⁸ to describe events. Detailed international biostratigraphic, palaeoecological and sedimentological study of many sections is required to determine whether transgressive pulses or depth maxima of local sequences correlate exactly. The former are not expected to do so, being reliant on local hypsographic curves, but the latter should correlate if they result from eustasy.

Daleje event

The first coiled Ammonoidea appeared in the earliest Emsian (Fig. 3a), possibly associated with an early Zlichovian transgressive event; the group became widespread globally within this period. The most distinctive goniatites were the Mimosphinctidae, their shells showing loosely-coiled spires with an

open, perforate umbilicus and whorls touching only in the outer stages of later members. Also present in the late Zlichovian were an aberrant group, the Auguritidae, and the Mimoceratidae (represented solely by *Gyroceratites*). In this earliest packet, therefore, goniatites ranged from early very simply sutured forms to more complex types (Fig. 4a).

In Czechoslovakia, the main group of mimosphinctids and the auguritids became extinct near the end of the Zlichovian⁹. Some mimosphinctids and *Gyroceratites* lingered into the overlying Dalejan, representing the diversity low. Then, in the higher Dalejan, two stocks significant in the later middle Devonian appeared, the Agoniatitidae and the Anarcestaceae. Their shell-forms were characterized by distinctive sutural patterns achieved early in ontogeny and soon achieved the imperforate umbilicus which characterized later Ammonoidea, to be followed by increased sutural complexity (Fig. 4b).

In the Czech sequence there is a striking correlation between the Daleje extinction event and the facies deepening seen in the early part of the Daleje Shales. A similar relation is seen in the transition from the Dede to Gebze formations of Bithynia Turkey¹⁰, and above the Sili'ninski Horizon of Novaya Zemlya¹¹. In the Khodza Kurgan Gorge of the Tien Shan, the faunal break lies in the Dzhaus Beds¹². The early Zlichovian transgression corresponds to the *Favosites regularissimus* Zone, here, in the Urals and elsewhere, and is succeeded by the Daleje event. In China the faunal¹³ and facies¹⁴ change is represented by the introduction of black mudstones in the lower Napiao Formation of Nandan, Kwangsi. It has been suggested⁸ that the Zlichovian event corresponds to the Esopus deepening in New York (Fig. 2) but the ammonoid evidence is not conclusive.

Kačák event

This next significant event in goniatite history has been used to define the base of the Maenioceras Stufe (Fig. 3a). A progressive decline in anarcestid diversity and, strikingly, the disappearance of the wide-ranging *Pinacites* preceded the event. The Tornoceratina then appeared; this group was characterized by a novel mode of sutural ontogeny with the A-type method of adult lateral lobe insertion¹. *Maenioceras* appeared and *Sobolewia* (traditionally grouped with the Prolobitidae) arose also at about this time¹; this genus was characterized by shells of convex (rather than biconvex) growth lines and globular form. The new fauna and a facies deepening are shown by the Kačák Member of the Sbskro Formation of Czechoslovakia⁹.

There was wide distribution of *Cabrieroceras*, especially the species *C. rouvillei*, *C. crispiforme* and *C. plebeiforme* (probably all synonyms) in Europe, North Africa and eastern North America¹⁵. With this distribution, together with evidence of other groups such as corals¹⁶, wide separation of Africa (south of the

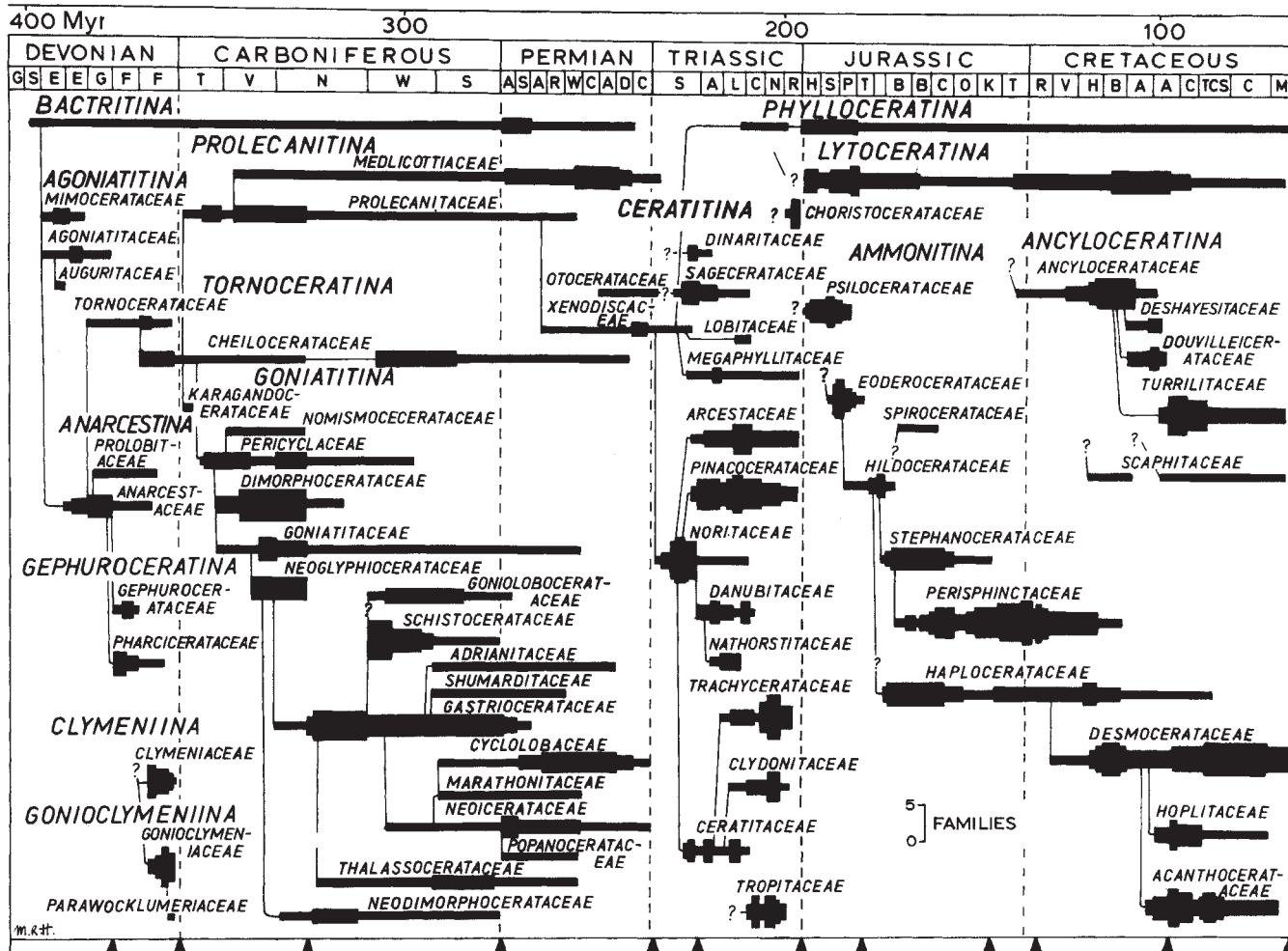


Fig. 1 Evolution of Ammonoidea based on data in ref. 1. Bactritina may originate in the Silurian and extend into the Triassic. Note the major extinction events at end of the Devonian, middle Carboniferous (Namurian E₂), end of the Permian, Triassic and Cretaceous. Such events have been used to define these major geological boundaries.

South Atlas Fault) from Europe and the Appalachians shown on some current palaeogeographical reconstructions¹⁷ seems unlikely. Progressive sutural complexity occurred also in the succeeding *Maenioceras* Stufe (Fig. 4c). There are other minor events in both New York⁸ and North Africa and Germany²⁵.

Taghanic event

The top of the *Maenioceras* Stufe is marked by the extinction of the Agoniatitidae, Pinacitidae and most of the Anarcestidae. Other than *Tornoceras* and possibly *Protornoceras*, no goniatite genus crossed this boundary. Sutural ontogeny of the subsequent novel groups indicates that they were descended from *Maenioceras* or anarcestids; only weak evidence exists that decline preceded the event.

In North America the deepening event was the Taghanic Onlap^{18,19}, or the Upper Devonian transgression of the Old World¹⁹, a term made redundant by recent unfortunate recommendations of the Devonian Subcommission²⁰. The event was initiated in the Middle *varcus* Zone.

This extinction event was significant for other groups also. Among brachiopods, about six families were lost^{21,22} and *Stringocephalus*, a common and worldwide marker for the Givetian, virtually disappeared, with some 15 coral families^{21,23}. These events imply significant modification of the widespread earlier Givetian carbonate platforms and its benthic environment as a result of the transgression, but this and other extinctions must be re-examined in the light of the precise zonations now available and crude studies based on stages abandoned.

The newly-arising goniatite stock was the Pharciceratidae, characterized by proliferation of umbilical lobes. The

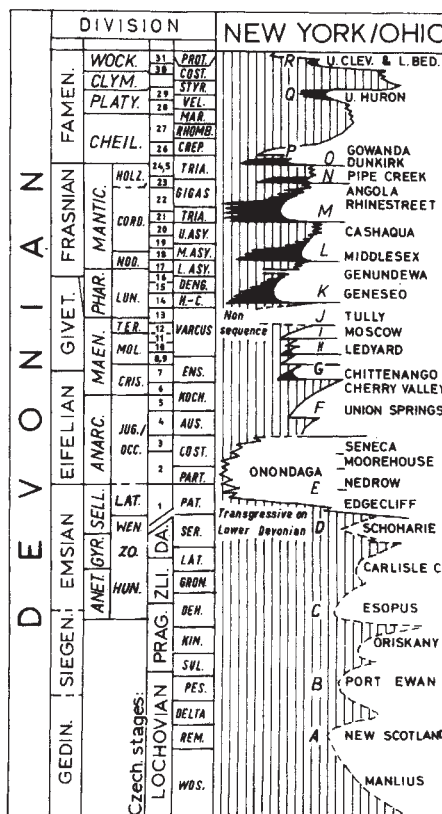


Fig. 2 New York Devonian succession illustrating facies movements and goniatite and conodont zonation⁸. Onlap movements move to the left (west); main black shale events are shown in black; numbers refer to local ammonoid zones¹⁵; standard stage, ammonoid and conodont zones are indicated.

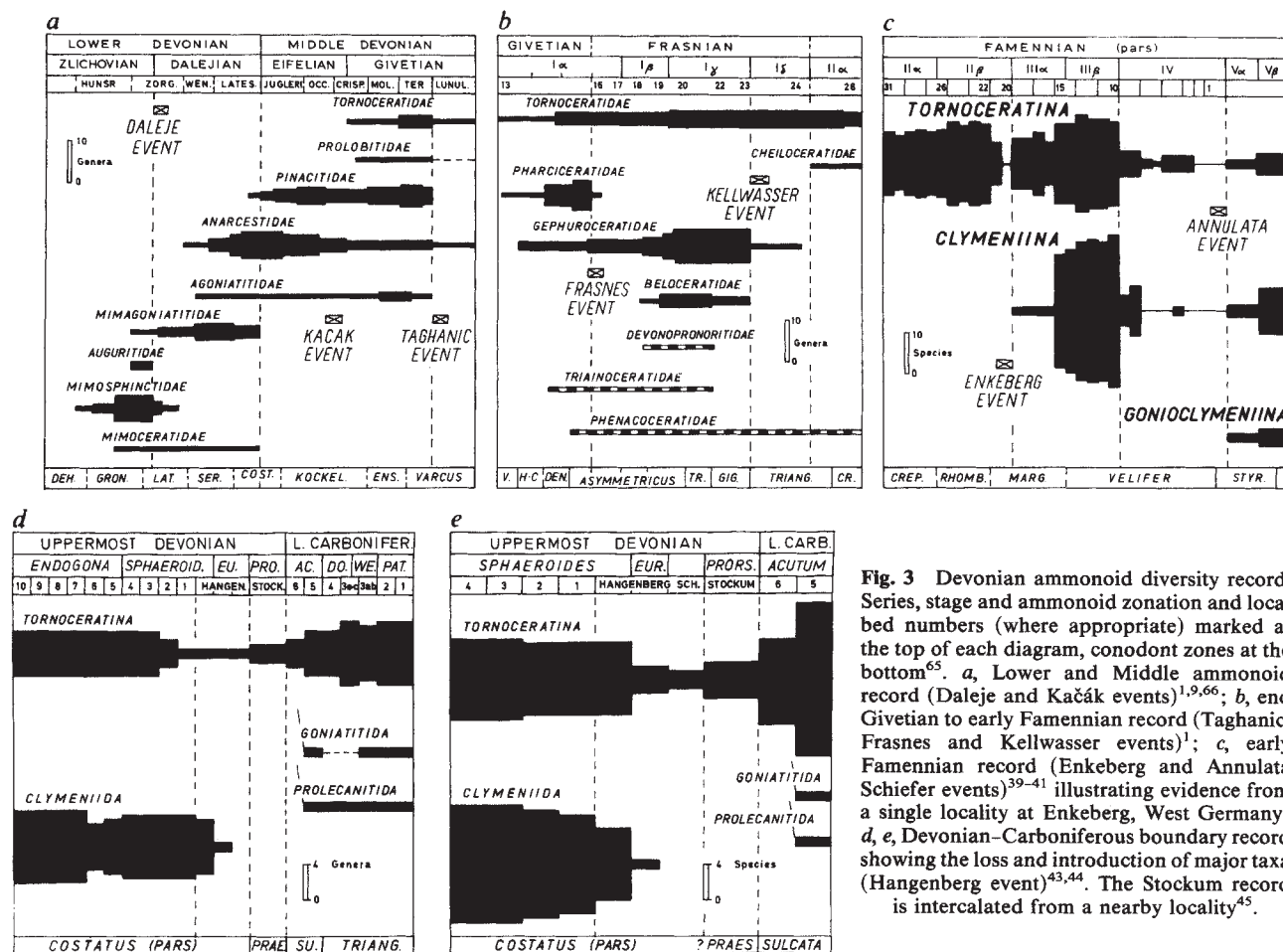


Fig. 3 Devonian ammonoid diversity record. Series, stage and ammonoid zonation and local bed numbers (where appropriate) marked at the top of each diagram, conodont zones at the bottom⁶⁵. **a**, Lower and Middle Devonian record (Daleje and Kacák events)^{1,9,66}; **b**, end Givetian to early Famennian record (Taghanic, Frasnian and Kellwasser events)¹; **c**, early Famennian record (Enkeberg and Annulata Schiefer events)³⁹⁻⁴¹ illustrating evidence from a single locality at Enkeberg, West Germany; **d**, **e**, Devonian-Carboniferous boundary record showing the loss and introduction of major taxa (Hangenberg event)^{43,44}. The Stockum record is intercalated from a nearby locality⁴⁵.

Gephuroceratidae, with a magnosellar suture, appeared shortly afterwards, as did a distinctive tornoceratid, *Epitornoceras*. In the Pharciceratidae, a progressive increase in sutural complexity occurred (Fig. 4d). The interval dominated by this family, the *lunulicosta* Zone, is divisible into five true zones, corresponding to Oberdevon to Ia of German nomenclature and approximately F₁ of classic Belgian terminology. It is renamed here the Pharciceras Stufe, defined at its base by the entry of *Pharciceras* in New York²⁴ and at the top by the entry of *Manticoceras* in the Upper Genundewa²⁴. The *Manticoceras* Stufe is thus returned to its original definition. The new base for the Upper Devonian currently proposed by the Devonian Subcommittee²⁰ (the base of the lower *asymmetricus* Zone) then falls only fractionally below the Pharciceras Stufe-Manticoceras Stufe boundary.

The interval of the Pharciceras Stufe which comprises the packet characterized by the Pharciceratidae corresponds to a period of radiation of palatolepid conodonts²⁵ and the rise in biomass of planktonic Ostracoda²⁶. There is little evidence of change in other groups; this is one reason why it seems unwise to use it to define a major series boundary²⁴.

Frasnes event

This break lies close to the base of the Lower *asymmetricus* Zone, corresponding approximately to the base of the Belgian Assise de Frasnian. Few pharciceratids survived the boundary; those that did produced the poorly-known Triainoceratidae and Phenacoceratidae (Fig. 3b). The best goniatite documentation is the New York section, where the upper Genundewa Limestone (Fig. 2; ref. 27) marks the entry of *Manticoceras*, a genus which soon became worldwide in its distribution (except for South America and Antarctica). Another new stock was the Beloceratidae (Figs 3b, 4e). In New York the Genundewa rep-

resents a transgressive level which may be correlated with the underlying remainie bed in western New York State. In Morocco, a good quantity of pharciceratid faunas continues up to the immediately underlying Lowermost *asymmetricus* Zone, but then is lost suddenly; the Lower *asymmetricus* Zone is represented, for example at Bou Tcharafine^{28,29}, by *Styliolina* shales, suggesting a deepening pulse. In Europe, this event marked the end of platform reefs. In this packet there is evidence also of progressive size increase and sutural complexity in *Manticoceras* and the Beloceratidae (Fig. 4e).

Kellwasser event

The term is used here to describe the Frasnian-Famennian faunal crisis^{19,30}, the event preceding any historic use of the boundary. The last of the Beloceratidae occurred in the latest Upper *gigas* Zone. At about the same level there was a decline in the Gephuroceratidae, leading to loss of all genera except *Crickites* and perhaps *Manticoceras*. In New York this occurred by the early Hanover Shale²⁷; in Belgium it fell in the Schistes de Matagne above *gigas* Zone levels. In the Timan the faunal break occurred in the Mendym, above the main euxinic oil-bearing shales of the Domanik³¹.

Following this event there were no major new ammonoid groups until the base of the Cheiloceras Stufe (Fig. 3b), when the Cheilocerataceae appeared with tornoceratid-type sutural ontogeny but convex growth lines. Two events rather than one may have occurred in the late Frasnian, represented by the end *gigas* Zone and the later loss of *Crickites* in the *triangularis* Zone. In other groups there was a sharp break in the conodont record at the close of the *gigas* Zone³², a large change in the coral record³³ and loss of dacryoconarids and many brachiopod families^{19,21,30}.

In the late Frasnian of Germany, a series of black limestone

pulses culminates in the main Kellwasserkalk of the top Upper *gigas* Zone at the time of most last appearances. The transgressive and euxinic nature of the Kellwasserkalk levels has been documented in detail³⁴ and a similar event occurred widely in Europe and North Africa³². Detailed work is still needed on the New York sequence, which contained pulses of black shales culminating in the thick Rhinestreet Shale in New York state (Figs 2, 5a), although the Pipe Creek Shale may have closer time analogy with the main Kellwasserkalk. This level represented the acme of global Devonian marine transgressions.

Frasnian extinctions can be correlated with evidence of marked environmental change in most known sections. The New York facies shifts (Figs 2, 5a) indicate a series of shallowing upward major rhythms⁸. A similar discontinuous pattern occurs in southern Belgium (Fig. 5b), northern France^{33,35} and western Canada (Fig. 5c)³⁶. Western Australian reefs³⁷ show a rather similar pattern and an abrupt termination by drowning of the stromatoporoid reef complex.

Precise global dating of these movements is needed to test for their synchrony. Conodont and ammonoid dating are both incomplete. For example, the sudden periods of Belgian and Canadian reef formation may correspond only to the initial times of black shale deepening in New York, even if these were caused by global eustatic events. Local depth curves are thus not directly comparable, and periods of still stand could have occurred for some time after reef termination (documented for the Canadian reefs³⁶).

Whether or not these events are synchronous, such major environmental changes must relate to the reduction of faunal diversity in the Frasnian and should therefore be taken into account in any hypothesis to explain it. The regressive events high in the *Cheiloceras* Stufe are too late to explain extinctions by 'perched faunas'³⁰. The virtual termination of reef complexes where stromatoporoids were frame builders was one of the striking features of late Frasnian palaeogeography and palaeoecology, and must be important in the loss of a large associated fauna³¹.

Enkeberg event

An event affecting ammonoid stocks at the closer of the Lower Famennian (Fig. 3c) conforms to the usual pattern, but is only well described from a locality at Enkeberg, West Germany³⁸⁻⁴¹. A fall in diversity occurred in the section culminating in Bed 20; this carried no fauna. Genera survived but *Cheiloceras* became extinct. Next, the Clymeniina appeared for the first time, its shell-form characterized by dorsal migration of the siphuncle in the earliest two whorls. Data from other areas confirm this significant faunal change in ammonoids, but no other section has been described in comparable detail.

Annulate event

The Annulate Schiefer of the German Famennian was a discrete event in the Rhenish Schiefergebirge. It consists of one (or a few) thin argillite unit(s) usually in pelagic carbonate sequences, where many costate platyclymenids are seen. This level marks the last appearance of true *Platyclymenia* and *Trigonoclymenia*⁴¹. The overlying Clymeniina Stufe is characterized by the entry of complex-sutured Gonioclymeniina. Simpler members of the Gonioclymeniina occur in the *Platyclymenia* Stufe, so that the suborder did not enter following the break (Fig. 3c). Generally, a significant European faunal change occurred above the break; genera such as *Cymaclymenia*, *Clymenia*, *Protoclymenia* and *Sellaclymenia* appeared⁴¹ and groups showed increasing sutural complexity (Fig. 4f). Detailed global correlation has not yet been attempted but platyclymeniid-bearing faunas are widespread globally (except South America and Antarctica). In the western United States, part of the Three Forks Shale has the same date⁴².

Hangenberg event

The base of the Carboniferous was defined by the Working Party on the Devonian-Carboniferous Boundary as the base of the

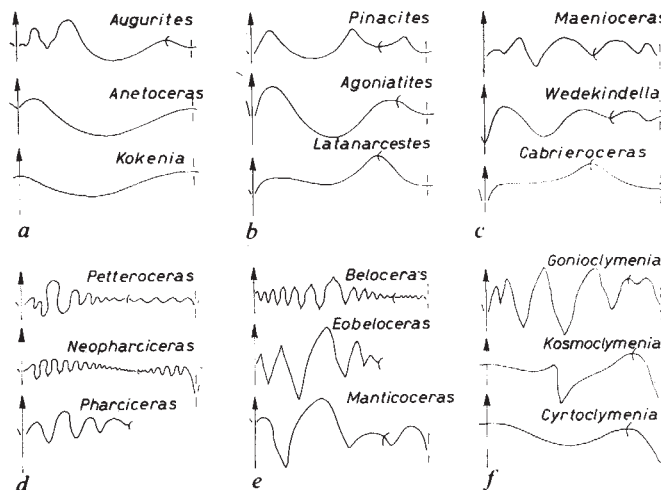


Fig. 4 Changes in ammonoid sutures showing increase in complexity between the extinction events discussed in the text. Exact evolutionary sequences not implied. a, Zlichovian; b, Dalejan and Eifelian; c, Givetian (early and mid); d, Givetian (late); e, Frasnian; f, Intra-Famennian.

conodont *sulcata* Zone. This corresponds closely to the Wocklumeria-Gattendorfia Stufe boundary recommended at the 1935 Heerlen Conference. The boundary ammonoid faunas at Oberödödinghausen, West Germany^{43,44} and the intercalated Stockum Limestone⁴⁵ are shown in Fig. 3d, e, the break marked by a decline and extinction of earlier stocks, followed by gradual increase in diversity and entry of two novel groups, the Prolecanitina (characterized by proliferation of U-type lobation and evolute form) and the Goniatitina (characterized by the division of the ventral lobe and A-type lobation). The extinction of other groups here has been documented^{31,46}.

The bleak period in the record corresponds to the euxinic black shale facies of the Hangenberg Schiefer. Estheriids occur in the shale at several localities and may indicate a near-shore influence. A more likely analogy for this phase, however, is the widespread occurrence of near-shore or terrestrial plant debris and fish remains in the Kellwasserkalk, expected from flooding over the alluvial plains, which may have carried debris out to sea⁴⁷. Similar facies changes occur in the Carnic Alps, Montagne Noire and Devon; the event is represented in North America by part of the Exshaw shale.

Tornocerataceae

In contrast with the history of other Devonian ammonoid stocks, the Tornocerataceae changed relatively little from their entry (at the Kačák event) until their extinction (at the Hangenberg event⁴⁸). Novel characteristics in this superfamily centred around shell-form modifications, particularly in the degree of umbilication, whorl outline, development of furrows and detailed growth line pattern. The elements of the suture varied very little and then only in relative proportions: few developed additional lobes or saddles and the basic A-type ontogeny remains unchanged¹. The Cheilocerataceae arose from the Tornocerataceae in the earliest Famennian; they differed significantly only in possessing convex growth lines. These two superfamilies belonged to the Tornoceratina, the only group of Devonian ammonoids recorded in South America⁴⁹. This is closest to the Devonian pole, suggesting that this group could survive in colder conditions than most ammonoids. The relations between the two superfamilies differed from other groups in the takeover pattern. The Tornocerataceae declined slowly, with a concurrent rise of the Cheilocerataceae (Fig. 6a), suggesting a more resilient life-style than other ammonoids. The Cheilocerataceae alone of ammonoid stocks survived the Hangenberg event and gave rise to later groups.

The Tornoceratidae⁵⁰ are among the few Palaeozoic ammonoid groups in which detailed evolutionary change has

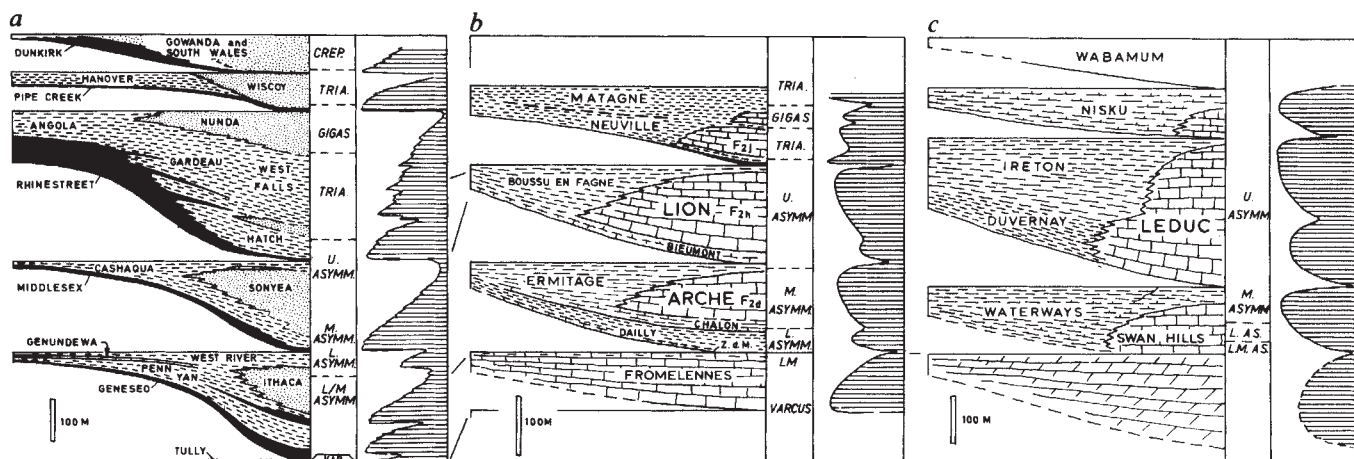


Fig. 5 Local successions and possible changes in sea level for Frasnian and adjacent rocks. *a*, New York State⁸; *b*, Belgium^{33,35}; *c*, western Canada^{37,67}. Black, black shales; dashes, shales; dots, siltstones and sandstones; 'bricks', limestones and dolomites. Onlap movements to the left.

been studied. Changes such as increasing amplitude of sutural element folding and increase in protoconch diameter occur in a Middle Devonian sequence, whereas shell-form shows no consistent change. Growth pattern could be defined statistically at particular faunal levels, but successive faunas behaved discontinuously and independently (Fig. 6*b*). The disjunct pattern may reflect a greater phenotypic control of shell-form than of other characteristics⁵⁰. If this is so, evidence of detailed evolutionary change in the fossil record will not be found without distinguishing between those characteristics of an organism which truly

mirror genetic change and those too strongly affected by environmental factors to be of value. It may not be possible, however, to identify true genetic change because of phenotypic masking. The distinct phenotypic forms which characterize particular horizons may reflect environmental changes in microrhythm event pulses of the type described in the Devonian⁸. Discrimination to 200–250 kyr was suggested for the Namurian, where goniatites occurred as part of sedimentary rhythm events⁵¹.

Conclusions

The evolutionary events described here may have a eustatic or euxinic immediate cause. There is, however, an unusual characteristic of some of the major sedimentary rhythms of the Devonian⁸, Carboniferous (Dinantian)⁵² and Jurassic^{8,51}; they are composed of microrhythms often sharing microfacies elements of the larger rhythms, for example, in the Lower West Falls Formation of New York⁸. It is difficult to explain this in wholly eustatic terms; a correlation of such microrhythms with Milankovitch-type cycles caused by orbital perturbations^{52–54} of some sort seems probable. Spread of cold conditions⁵⁵, CO₂ variations⁵⁶, associated changes in ocean currents, precipitation patterns and palaeogeographic modifications may have caused extinction events and phenotypic fluctuations; it is unlikely that the rhythm-associated sedimentology is simple.

It is difficult to determine whether the larger-scale evolutionary breaks described here are periodic. The eight extinction events span about 30 Myr (estimated from the Cambridge time scale⁵⁷), giving a hypothetical internal of ~4.3 Myr if all the events have been identified, which may not be the case. Radiometric data are too imprecise for use here, but field data suggest that the Taghanic and Frasnian events and the Enkeberg and Annulata events are closer together in time than others. Furthermore, three of the events have a major impact on other groups. Thus, there is a difference in degree as well as period.

The possibility of a periodic explanation of sedimentary and palaeontological events has been much discussed since rhythmicity was first observed and documented, even leading to astronomical causation theories^{58,59}. Recent suggestions include a solar companion star, the effects of meteorite impacts (bolides) and cosmic dust^{60,61}, alleging a 26-Myr extinction periodicity⁶⁰. But to which of the eight extinction events described here might this speculation apply? Three events affected other groups significantly (Taghanic, Kellwasser and Hangenberg). Yet with a 26-Myr periodicity, only one of these could be a contender and a different explanation would be needed for the remaining extinction events. However, all the events share the same six features described here. The most significant event may have been the Kellwasser³⁰; but the Taghanic event was associated

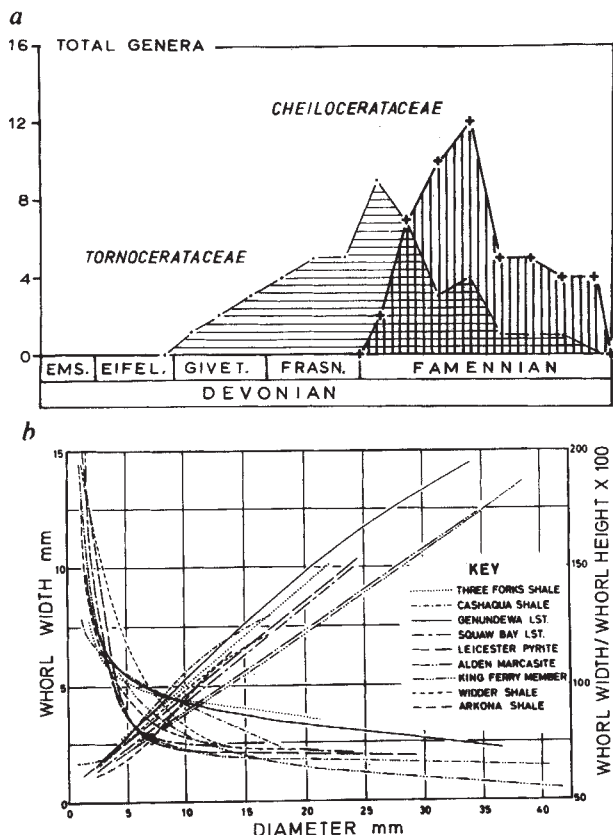


Fig. 6 *a*, Pattern of taxonomic diversity record in the Devonian Tornocerataceae and Cheilocerataceae showing progressive, rather than sudden, replacement. *b*, Inferred disparate phenotypic effects in shell-growth pattern from a sequence of *Tornoceras* from the Devonian of North America⁵⁰.

with the largest biomass change³¹. Recent geochemical work^{62,63} is ambiguous on whether iridium concentrations are associated with the Kellwasser event and hence whether cosmic showers may be important. But iridium data are needed from throughout sequences, not only from particular horizons. In the Devonian and Carboniferous^{63,64}, major sedimentary rhythms and extinction events did not occur at regular intervals nor were equal in effect. This suggests a more complex explanation than current astronomical theories provide.

During the eight events described here, the end Caledonian, Antler and Acadian orogenies and the Bretonic movements all occurred. These events will have had global eustatic effects to which sea-floor spreading rate perturbations will have contributed. Extraterrestrial influences (such as precession, tilt and eccentricity effects acting through solar insolation changes) would interact with terrestrial factors to give complex patterns; effects due to companion stars or asteroids might be added speculatively to this list. Progress will be made if the small-scale microrhythms, a common feature of the Phanerozoic record, can be correlated precisely with Milankovitch-type cycles to form the basis for a new and detailed chronology. This new objective time scale could revolutionize the study of Earth history.

Received 27 June; accepted 22 October 1984.

1. House, M. R. & Senior, J. R. (eds) *The Ammonoidea* (Academic, London, 1981).
2. Newell, N. D. *J. Paleont.* **36**, 371–385 (1952).
3. House, M. R. *Adm. Sci.* **19**, 155–163 (1962).
4. Ward, P. D. & Signor, P. W. III *Paleobiol.* **9**, 183–198 (1983).
5. Klüpfel, W. *Geol. Rdsch.* **7**, 97–109 (1916).
6. Ramsbottom, W. H. C. *J. geol. Soc.* **136**, 147–153 (1979).
7. Bayer, U. & McGhee, G. R. Jr, *Lethaea* **17**, 1–16 (1984).
8. House, M. R. *Proc. Usher Soc.* **5**, 396–405 (1983).
9. Chlupáč, I. & Turek, V. *Rozpr. ustr. Ust. Geol.* **66**, 1–159 (1983).
10. Hass, W. in *Biostratigrafiya Pogranichnykh Otlozhenii Nizhnego i Srednego Devonu* (eds Sokolov, B. S. & Rzhonsnitskaya, M. A.) 114–147 (Akad. Nauk SSSR, Leningrad, 1982).
11. Nekorosheva, L. V., Patrynov, D. K. & Sobolev, N. N. in *Biostratigrafiya Pogranichnykh Otlozhenii Nizhnego i Srednego Devonu* (eds Sokolov, B. S. & Rzhonsnitskaya, M. A.) 108–110 (Akad. Nauk SSSR, Leningrad, 1982).
12. Bogoslovskiy, B. I. in *Biostratigrafiya Pogranichnykh Otlozhenii Nizhnego i Srednego Devonu* (eds Sokolov, B. S. & Rzhonsnitskaya, M. A.) 23–26 (Akad. Nauk SSSR, Leningrad, 1982).
13. Ruan Yi-Ping, in *A Handbook of the Stratigraphy and Palaeontology in Southwest China* (in Chinese) (ed. Nanking Institute of Geology and Palaeontology) 238 (Academia Sinica, Peking, 1974).
14. Yang, S., Pan, K. & Hou, H. *Geol. Mag.* **118**, 113–224 (1981).
15. House, M. R. *Spec. Pap. Paleont. Soc.* **21**, 1–70 (1978).
16. Oliver, W. A. Jr, *J. Paleont.* **50**, 365–373 (1976).
17. Johnson, G. A. L. *Proc. Yorks. geol. Soc.* **43**, 221–252 (1981).
18. Johnson, J. G. *Bull. geol. Soc. Am.* **81**, 2077–2106 (1970).
19. House, M. R. *Proc. Yorks. geol. Soc.* **40**, 233–288 (1975).
20. Ziegler, W. & Klapper, G. in *Episodes* Vol. 4 (eds Agterberg, F. P., Gradstein, F. M. & Williamson, M.) 18–21 (1984).
21. Sepkoski, J. J. Jr, *Contr. biol. Geol. Milwaukee Publ. Mus.* **51**, 1–125 (1982).
22. Boucot, A. *Evolution and Extinction Rate Controls* (Elsevier, Amsterdam, 1975).
23. Oliver, W. A. Jr, & Pedder, A. E. H. *Spec. Pap. palaeont.* **23**, 233–248 (1979).
24. House, M. R. *Cour. Forsch.-Inst. Senckenberg* **55**, 449–462 (1982).
25. Helms, J. & Ziegler, W. in *Treatise on Invertebrate Paleontology Pt W* (ed. Robson, R. A.) W98–99 (Geol. Soc. Am. & University of Kansas Press, 1981).
26. Rabien, A. & Rabitz, A. *Notizbl. hess. Landesamt. Bodenforsch. Wiesbaden* **86**, 159–184 (1958).
27. Kirchgasser, W. T. & House, M. R. in *Devonian Biostratigraphy of New York* (eds Oliver, W. A. Jr & Klapper, G.) 39–55. (IUGS Subcommittee on Devonian Stratigraphy, Washington, 1981).
28. Bultynck, P. & Hollard, H. *Aardkundige Meded.* **1**, 1–73 (1980).
29. Bensaid, M., Bultynck, P., Walliser, O. H. & Ziegler, W. *Proposal of a Type Section for the Middle/Upper Devonian Boundary in Morocco* (IUGS, Devonian Subcommittee, Washington, 1983).
30. McLaren, D. J. *Bull. geol. Soc. Am.* **190**, 477–485 (1982).
31. House, M. R. *Proc. Yorks. geol. Soc.* **40**, 459–490 (1975).
32. Büggisch, W. & Clausen, C.-D. *Neues Jb. Geol. Paläont. Abh.* **141**, 137–167 (1972).
33. Tsien, H. H. in *Guidebook 2nd int. Symp. Fossil Corals and Reefs. Excursion C*, 3–43 (Serv. geol. Belgique, Bruxelles, 1975).
34. Büggisch, W. *Abh. hess. Landesamt. Bodenforsch.* **62**, 1–62 (1972).
35. Mouraviev, A. N. in *Pap. Frasnian Givetian Boundary* (eds Bigey, F. et al.) 101–118 (Serv. geol. Belgique, Bruxelles, 1982).
36. Playford, P. E. *Bull. Am. Ass. Petrol. Geol.* **64**, 814–840 (1980).
37. Mountjoy, E. W. in *Int. Symp. Devonian System*, Calgary, Vol. 2 (ed. Oswald, D. H.) 387–418 (Alberta Soc. Petrol. Geol., 1968).
38. Ziegler, W. *Abh. Hess. Landesamt. Bodenforsch.* **38**, 1–166 (1962).
39. Wedekind, R. *Neues Jb. Geol. Paläont. Abh.* **26**, 565–634 (1908).
40. Lange, W. *Abh. preuss. geol. Landesanst.* **119**, 1–132 (1929).
41. Price, J. D. thesis, Univ. Hull (1983).
42. House, M. R. *J. Paleont.* **36**, 247–284 (1962).
43. Schindewolf, O. H. *Abh. preuss. geol. Landesanst.* **178**, 1–132 (1937).
44. Vöhringer, E. *Fortschr. Geol. Rheinl. Westf.* **3**, 107–197 (1960).
45. Alberti, H., Groos-Uffenorde, H., Streel, M., Uffenorde, H. & Walliser, O. H. *Newslett. Stratigr.* **3**, 263–276 (1974).
46. House, M. R., Scrutton, C. T. & Bassett, M. G. (eds) *The Devonian System* (Palaeontology Association, London, 1979).
47. Bogoslovskiy, B. I. *Trudy paleont. Inst.* **124**, 1–341 (1969).
48. House, M. R. & Price, J. D. *Palaeontology* **28**, 159–188 (1985).
49. Hünicken, M., Kullmann, J. & Suarez Riglos, M. *Bol. Acad. nac. Ciencias Cordoba* **53**, 238–253 (1980).
50. House, M. R. *Phil. Trans. R. Soc.* **B250**, 79–130 (1965).
51. House, M. R. *The Dorset Coast from Poole to the Chesil Beach* (Geological Association, London, 1969).
52. Schwarzscher, W. in *Episodes* (eds Agterberg, F. P., Gradstein, F. M. & Williamson, M.) Vol. 7(1), 64 (1984).
53. Milankovitch, M. *Handbuch der Klimatologie* Vol. 1A (Bornträger, Berlin, 1930).
54. Imbrie, J. & Imbrie, J. Z. *Science* **207**, 943–953 (1980).
55. Copper, P. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **21**, 165–207 (1977).
56. Pisias, N. G. & Shackleton, N. J. *Nature* **310**, 757–759 (1984).
57. Harland, W. B. et al. *A Geologic Time Scale* (Cambridge University Press, 1982).
58. Adhemer, A. J. *Revolutions de la Mer* (Carilian Gœury & V. Dalmont, Paris, 1842).
59. Croll, J. *Climate and time in their Geological Relations, a Theory of Secular Change of the Earth's Climate* (Daldy, Ibister & Co., London, 1875).
60. Raup, D. M. & Sepkoski, J. J. Jr, *Proc. natn. Acad. Sci. U.S.A.* **81**, 801–805 (1984).
61. Kerr, R. A. *Science* **223**, 1277–1279 (1984).
62. McGhee, G. R., Jr, Golmore, J. S., Orth, C. J. & Olsen, E. *Nature* **308**, 629–631 (1984).
63. Ramsbottom, W. H. C. *Proc. Yorks. geol. Soc.* **14**, 261–291 (1977).
64. Kullmann, J. *Paläont. Z.* **57**, 231–240 (1983).
65. Klapper, G. & Ziegler, W. *Spec. Pap. palaeont. Soc.* **23**, 199–224 (1979).
66. Bensaid, M. *Notes Serv. geol. Maroc* **36**, 81–140 (1974).
67. Norris, A. W. & Uyeno, T. T. *Bull. geol. Surv. Can.* **313**, 1–65 (1983).
68. Playford, P. E., McLaren, D. J., Orth, C. J., Gilmore, J. S. & Goodfellow, W. D. *Science* **226**, 437–439 (1984).

RNA polymerase heterogeneity in *Streptomyces coelicolor*

Janet Westpheling, Monica Raney & Richard Losick

Department of Cellular and Developmental Biology, Harvard University, 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA

Two forms of RNA polymerase holoenzyme have been identified in the filamentous differentiating bacterium Streptomyces coelicolor. They contain different species of σ factor and are distinguishable by their ability to recognize different promoter classes. These and other holoenzyme forms may in part determine the selective expression of different gene sets in this morphologically-complex bacterium.

BACTERIA of the Gram-positive genus *Streptomyces* are unusual among prokaryotes in the complexity of their morphological differentiation cycle and the large number and diversity of antibiotic compounds produced as secondary metabolites, many of which are medically and agriculturally important¹. *Streptomyces* grow as a complex substrate mycelium with branching hyphae penetrating the soil and solubilizing organic material

by the secretion of extracellular hydrolytic enzymes. Under conditions of nutrient limitation, the bacteria differentiate by the formation of specialized aerial hyphae which coil and then septate to form chains of spores. The production of antibiotics and other secondary metabolites often coincides with the onset of morphological development, and thus may be regulated by the same mechanisms as those that trigger differentiation¹. The

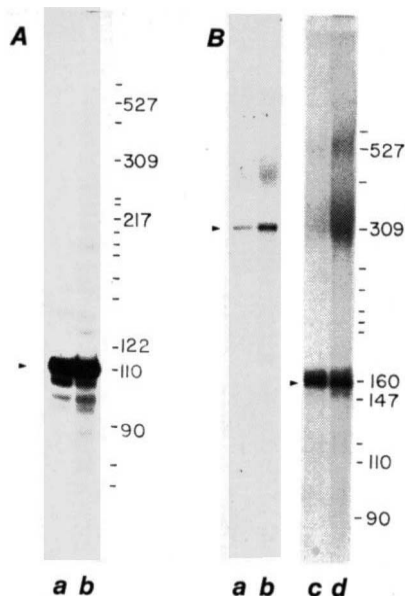


Fig. 1 Run-off transcription by RNA polymerase from *S. coelicolor* and *B. subtilis*. **A**, Run-off RNAs from the *veg* promoter contained in *Bam*HI-cut pMS500 DNA were generated by *B. subtilis* $E\sigma^{55}$ (lane *a*) or by *S. coelicolor* RNA polymerase (lane *b*). **B**, Run-off RNAs from the *ctc* promoter were generated by *B. subtilis* $E\sigma^{37}$ using *Pst*I-cut pMI520 DNA (lane *a*) or *Bam*HI-cut pMI340 DNA as templates (lane *c*) or by *S. coelicolor* RNA polymerase using *Pst*I-cut pMI520 DNA (lane *b*) or *Bam*HI-cut pMI340 DNA (lane *d*) as templates.

Methods: *S. coelicolor* A3(2) cells were grown in YEME medium (0.3% Difco Bacto yeast extract, 0.5% Difco Bacto peptone, 0.3% Oxoid malt extract, 0.1% $MgCl_2$, 1% glucose, 34% sucrose) and collected in early stationary phase. A crude preparation of RNA polymerase was prepared from 100 g cells suspended in buffer A (0.01 M Tris pH 8.0, 1 mM EDTA, 0.3 mM dithiothreitol, 0.3 mg ml⁻¹ phenylmethylsulphonyl fluoride, 10% glycerol) containing 0.1 M KCl by passage of the suspended cells through a French pressure cell (10,000 p.s.i.) and centrifugation for 30 min at 100,000g. RNA polymerase was partially purified by applying the supernatant to a heparin-agarose column (1.3 × 5 cm), washing the column with buffer A containing 0.1 M KCl, then eluting in one step with 3 ml buffer A containing 0.5 M KCl. 10 µl of the eluate (0.0004 units of poly(dA-dT) transcribing activity as defined in ref. 40) or 0.001 units of the indicated *B. subtilis* holoenzyme were incubated at 37 °C for 10 min in 40 µl reactions⁷ with 2 µg of the indicated template. RNA synthesis was initiated by the addition of 0.5 mM each of ATP, GTP and UTP and 0.5 µM (10 µCi [³²P]) CTP. After 1 min at 37 °C, 6 µg heparin was added to prevent reinitiation and, after 5 min, 1.5 mM unlabelled CTP was added. After a further 5 min, the reaction was stopped on ice. Radioactive RNA was concentrated from the reaction mixtures by ethanol precipitation and subjected to electrophoresis in a 6% polyacrylamide gel containing 7 M urea. Molecular weight markers were heat-denatured ³²P-end-labelled fragments of *Hpa*II-cut pBR322 DNA. pMS500 has been described⁷. pMI340 and pMI520 were constructed by M. Igo (unpublished results), containing *Bam*HI sites and *Pst*I sites at 155 and 261 bp, respectively, downstream from the *ctc* transcription initiation site.

regulation of developmental genes and genes involved in the production of antibiotics have been studied², although little is yet known about the mechanisms governing the transcription of these and other kinds of complex gene sets.

Developmental gene regulation is better understood in *Bacillus*, another group of Gram-positive, spore-forming bacteria (reviewed in refs 3, 4). An important feature of the transcriptional apparatus of these bacteria relevant to the expression of certain complex gene sets (including developmental genes), is the existence of multiple forms of RNA polymerase holoenzyme, each with the capacity to use a characteristic class of promoters^{5,6}. Each of these holoenzyme forms is associated with one of several species of σ factor, which confers on a common-

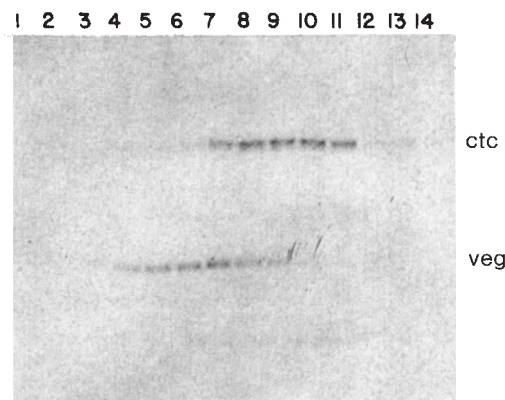


Fig. 2 Partial separation of *veg*- and *ctc*-transcribing activities. Run-off RNAs were generated from a mixture of *veg*- and *ctc*-containing templates by *S. coelicolor* RNA polymerase in fractions from a DNA cellulose gradient elution. The positions of authentic *veg* and *ctc* run-off RNAs generated by $E\sigma^{55}$ and $E\sigma^{37}$ from *B. subtilis* are indicated (right).

Methods: *S. coelicolor* A3(2) cells (350 g) grown in YEME medium to early stationary phase were disrupted in a French press and RNA polymerase partially purified by phase partitioning and gel filtration as described previously⁴⁰, except that size fractionation was on a Sephacryl S300 column (3 × 40 cm). Fractions containing the peak of RNA polymerase activity (5,300 units) determined by incorporation of ³H-UTP into RNA with poly(dA-dT) as template⁴⁰ (26 units per mg protein) were pooled and purified by gradient elution from heparin-agarose (1.3 × 5 cm) twice in succession and then from a DNA-cellulose column. The first heparin-agarose column was eluted with a 0.05–0.6 M linear KCl gradient and the second eluted with a 0.25–0.53 M linear KCl gradient. The first DNA-cellulose column was eluted with a 0.15–0.5 M linear gradient of KCl. In the final purification step, RNA polymerase in fractions containing *veg*- and *ctc*-transcribing activities were reappplied to DNA-cellulose (a 0.5 × 5 cm column) and eluted with a linear KCl gradient of 0.3–0.45 M KCl in 30 ml buffer A. Fractions of 3 ml were collected (fraction 1 was from the low end of the salt gradient) and 20 µl samples assayed by run-off transcription as described in Fig. 1, using a mixture of *veg* (*Bam*HI-cut pMS500)- and *ctc* (*Bam*HI-cut pMI340)-containing templates. The ³²P-labelled products of RNA synthesis were displayed by electrophoresis in a 6% polyacrylamide gel containing 7 M urea and visualized by autoradiography.

core RNA polymerase moiety the ability to respond to a particular set of transcription initiation signals. We seek here to elucidate whether transcription from different gene sets in *Streptomyces* is determined, in part, by multiple forms of RNA polymerase, and if so, whether some of these holoenzyme forms might be homologous in their promoter recognition specificity to previously-described forms of RNA polymerase from *B. subtilis*, the most extensively studied of the *Bacillus* species. We report the discovery of two distinct forms of RNA polymerase holoenzyme from *S. coelicolor* through the use of well-characterized promoters from *B. subtilis* as templates for *in vitro* transcription and the identification of a *bona fide Streptomyces* promoter whose use is determined by one of these holoenzyme forms.

S. coelicolor RNA polymerase

To examine promoter recognition specificity of *S. coelicolor* RNA polymerase, we used plasmid DNAs containing either of two well-characterized promoters from *B. subtilis*, *veg* and *ctc*, as templates for *in vitro* transcription. The *veg* promoter is used by the predominant form of RNA polymerase in *B. subtilis*, $E\sigma^{55}$ (σ subunit of relative molecular mass (M_r) 55,000) which recognizes *B. subtilis* promoters that, like *veg*, strongly conform to the consensus *Escherichia coli* promoter sequence prototypical for eubacteria⁷. The *ctc* promoter, however, is not recognized by $E\sigma^{55}$, but is used by minor forms of *B. subtilis* RNA polymerase, including in particular the $E\sigma^{37}$ form (σ subunit of M_r 37,000), whose interaction with the *ctc* promoter has been studied in detail^{8–11}.

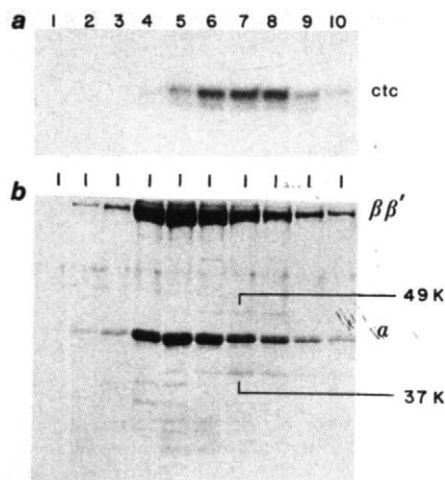


Fig. 3 Subunit composition of the *ctc*-transcribing form of *S. coelicolor* RNA polymerase. *a*, Run-off RNAs were generated from the *ctc* promoter contained in *Bam*HI-cut pMI340 by purified RNA polymerase in fractions from a DNA-cellulose gradient elution. *b*, Proteins contained in fractions from the gradient elution were subjected to electrophoresis in a 10% SDS polyacrylamide gel and then stained with Coomassie brilliant blue.

To detect transcription initiation from the *veg* and *ctc* promoters, *S. coelicolor* RNA polymerase, partially purified from a cleared cell lysate by stepwise elution from heparin-agarose, was used to generate 'run-off' transcripts with linearized plasmid templates containing the *veg* or *ctc* promoters. *S. coelicolor* RNA polymerase generated a run-off RNA (Fig. 1A, lane *a*) indistinguishable in size from that produced by the *B. subtilis* Eo⁵⁵ enzyme (Fig. 1A, lane *b*) with linearized *veg*-containing DNA as template. Similarly, the run-off RNAs copied from two *ctc*-containing templates by the *S. coelicolor* enzyme (Fig. 1B, lanes *a* and *c*) were apparently identical to the corresponding RNAs produced by *B. subtilis* Eo³⁷ enzyme (Fig. 1B, lanes *b* and *d*). Thus, a relatively crude preparation of RNA polymerase from *S. coelicolor* accurately uses both heterologous promoters, initiating at sites close to or identical with those from which the corresponding *B. subtilis* enzymes initiate transcription.

We next purified further the *veg*- and *ctc*-transcribing activities (see legend to Fig. 2). Gradient elution partially resolved *veg*-transcribing activity from *ctc* transcribing activity (Fig. 2), indicating that the *veg* and *ctc* promoters are separately used by different forms of *S. coelicolor* RNA polymerase.

To analyse the subunit composition of the enzyme responsible for *ctc* transcription, this form of RNA polymerase was purified further (see Fig. 3 legend). Gradient-purified RNA polymerase contained, in addition to several unidentified polypeptides, core subunits β' , β and α (as observed previously for RNA polymerase from *S. antibioticus*¹² and *S. coelicolor*¹³) and two polypeptide species of $M_r \sim 49,000$ and $37,000$ whose elution profile coincided closely with that of the *ctc*-transcribing activity (Fig. 3).

Two species of RNA polymerase σ factor

We next sought to determine whether either of these polypeptides were responsible for the *ctc*-transcribing activity of *S. coelicolor* RNA polymerase, and if so, whether such a factor(s) could be considered functionally equivalent to one of the various σ species associated with *B. subtilis* RNA polymerase. To investigate this possibility, we performed a heterologous complementation experiment in which RNA polymerase-associated polypeptides from *S. coelicolor* were tested in the presence of core RNA polymerase from *B. subtilis* for their ability to stimulate transcription from the *ctc* promoter. RNA polymerase-associated polypeptides were isolated as described previously¹⁴. We found that a protein of $M_r \sim 49,000$ strongly stimulated the transcription of *ctc* DNA by core RNA polymerase (Fig. 4A, C).

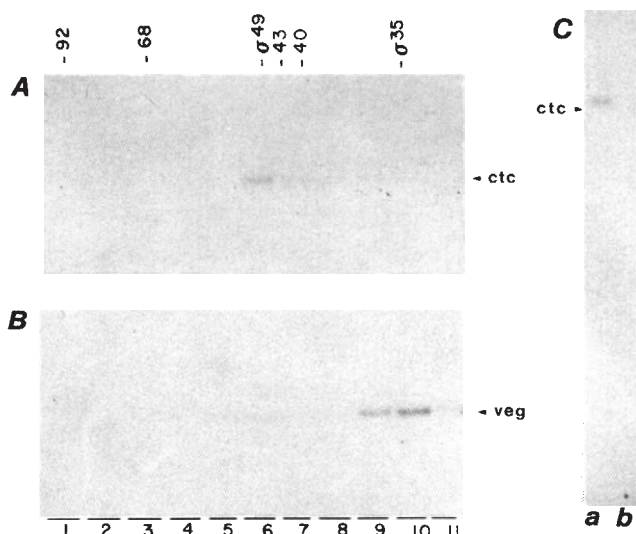


Fig. 4 Complementation of *B. subtilis* core RNA polymerase with gel-purified RNA polymerase-associated polypeptides from *S. coelicolor*. *A*, *B*, Parts of *ctc* and *veg* run-off RNA electropherograms generated by *B. subtilis* core RNA polymerase supplemented with gel-purified proteins; only the size range of *ctc* and *veg* transcripts is shown. *C*, Full-length electropherogram of run-off RNA generated by *B. subtilis* core RNA polymerase supplemented with protein from slice 6 (lane *a*) and run-off RNA generated by *B. subtilis* core enzyme alone (lane *b*).

Methods: RNA polymerase was partially purified from 300 g *S. coelicolor* cells through the first heparin-agarose elution step of the procedure in Fig. 2 legend. (At this stage, the enzyme was relatively impure, containing both *veg*- and *ctc*-transcribing activities.) RNA polymerase ($\sim 25 \mu\text{g}$ protein) in fractions containing *veg* and *ctc* transcribing activity were pooled and subjected to electrophoresis in a 12% (Laemmli) polyacrylamide gel, stained briefly with Coomassie brilliant blue. The gel was then sliced from top to bottom into 11 slices; protein from each of the slices was eluted and renatured by the procedure of Hager and Burgess¹⁴. $20 \mu\text{l}$ (3.6% of the total volume of renatured protein) were incubated with 0.001 units of *B. subtilis* core RNA polymerase at 37°C for 5 min with $25 \mu\text{l}$ transcription reaction mixture⁴⁰. DNA template ($2 \mu\text{g}$) was then added and incubated for 3 min to allow prebinding of polymerase. RNA synthesis was then initiated as described in Fig. 1 legend.

In other experiments, we have observed also significant stimulatory activity at the position of the $M_r 37,000$ protein (J. Guijarro and J.W., unpublished observations). The relative proportion of the $M_r 49,000$ and $37,000$ stimulating activities varies significantly from experiment to experiment; we do not know whether the $M_r 37,000$ species is a proteolytic fragment of the $M_r 49,000$ polypeptide or an additional *ctc* stimulatory factor. We conclude that at least the $M_r 49,000$ polypeptide (σ^{49}) and also a $M_r 37,000$ species (σ^{37}) are σ -like transcription factors that dictate recognition of the *ctc* promoter.

To identify a corresponding transcription factor that allows use of the *veg* promoter, renatured proteins were tested for their ability to stimulate transcription of *veg*-containing template by *B. subtilis* core enzyme. A polypeptide of $M_r \sim 35,000$ (but clearly smaller than the $M_r 37,000$ *ctc* stimulatory factor), σ^{35} , strongly enhanced *veg* RNA synthesis above a low background of transcription by the core enzyme preparation alone (Fig. 4B).

Transcription

We have identified a *Streptomyces* promoter used by *ctc*-transcribing purified RNA polymerase, but not *veg*-transcribing enzyme. This promoter directs the transcription of a *S. plicatus* gene, *endoH*^{15,16}, whose product is a secreted endo- β -N-acetylglucosaminidase (endoglycosidase-H) that hydrolyses glycosidic bonds between the *N*-acetylglucosamine residues of 'high mannose' oligosaccharides¹⁷. We have used the cloned *endoH* gene¹⁵ as a hybridization probe for identifying the *endoH*

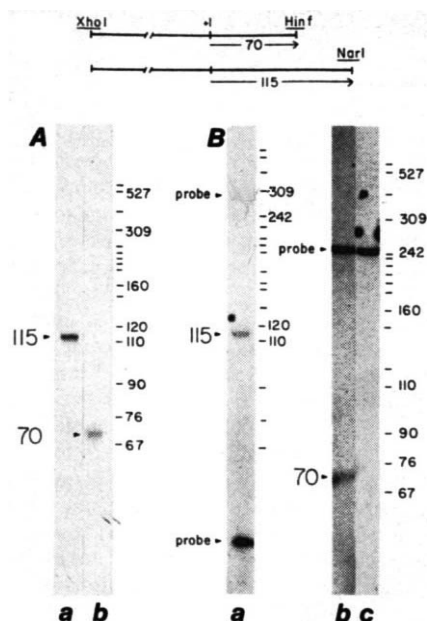


Fig. 5 Identification of the transcription start site of the *S. plucatus* *endoH* promoter and its utilization by *S. coelicolor* RNA polymerase *in vitro*. Top, DNA fragments in the region just upstream to the *endoH* coding sequence, gel-purified by electrophoresis of restriction endonuclease-digested pEHB1.6 DNA¹⁵. A, *In vitro* run-off transcription from the *endoH* promoter. Run-off RNAs were generated by the *ctc*-transcribing form of *S. coelicolor* RNA polymerase (purified as described in the Fig. 3 legend) from *XhoI*-*NarI* (lane a) and *XhoI*-*HinfI* (lane b) DNA fragments of 300 and 265 bp, respectively. B, S₁ nuclease mapping of *endoH* RNA. The 5' end of *endoH* RNA from *S. plucatus* cells was mapped by using *XhoI*-*NarI* DNA (lane a) and *XhoI*-*HinfI* (lane b) DNA as radioactive probes. Lane c, a control in which *XhoI*-*HinfI* probe was annealed with *E. coli* RNA.

Methods: B, 100 ml *S. plucatus* cells were grown in YEME medium for 6 days and RNA extracted as described previously⁴¹ except that resuspension buffer contained 1% SDS (suggested by M. Igo). The DNA probes for S₁ mapping were prepared by digesting plasmid pEHB1.6 with either *HinfI* or *NarI*, then treating with calf intestine alkaline phosphatase. Next, the *HinfI* fragment was cut with *XhoI* to yield a 265-base *XhoI*-*HinfI* promoter-containing segment, then separated by gel electrophoresis from the remaining 1.4-kb fragment. The *NarI* fragment was cut simply with *XhoI* resulting in a 300-bp *XhoI*-*NarI* promoter-containing segment and smaller 30-bp fragment, and was not further purified. 100 µg RNA was mixed with ~0.7 µg (~3 × 10⁷ c.p.m. µg⁻¹) of ³²P-labelled DNA, and incubated at 85 °C for 15 min in hybridization buffer¹⁸. Hybridization was performed at 58–60 °C for 3 h and was quenched by the addition of 300 µl of dilution buffer, then treated with S₁ nuclease (400 units ml⁻¹) as described previously¹⁸. The S₁ nuclease-resistant hybrids were precipitated in ethanol at -70 °C and displayed by autoradiography after electrophoresis in 6% polyacrylamide gels containing 7 M urea.

```

TTTCGAGGTTTAAATCCTTATCGTTATGGGTAATTGTTGTAATAGG   cta
      ↑
ATTGACTGATTGACGCGC TTCCGGCGGGGCGAGGGAGGCACGGTG   endoH

```

Fig. 6 Nucleotide sequence of the *cta* and *endoH* promoter regions. Transcription start sites are underlined; other symbols are explained in the text. The sequence of the *endoH* promoter was determined by subjecting the end-labelled *XhoI*-*HinfI* and *XhoI*-*NarI* fragments described in Fig. 5 legend to the base-specific cleavage reactions of Maxam and Gilbert¹⁹.

transcription start site *in vivo* and as a template for reconstructing its specific transcription *in vitro*. The apparent start site was determined by S₁ nuclease experiments¹⁸, in which DNA fragments located upstream to the endoglycosidase-H coding sequence were hybridized to RNA isolated from *S. plucatus* cells. *S. plucatus* RNA protected a 115-base segment of a *XhoI*-*NarI* probe and a 70-base segment of a *XhoI*-*HinfI* probe from nuclease action (Fig. 5B, lanes a, b). Neither of these DNAs was observed in control experiments in which the end-labelled probes were incubated with RNA from *E. coli* cells (Fig. 5B, lane c; data not shown). These experiments identify the 5' terminus of *endoH* mRNA. This terminus was located approximately 140 base pairs (bp) upstream from the inferred *endoH* translation initiation codon (as identified previously from heterologous expression experiments in *E. coli* cells¹⁶). A more precise location of the *endoH* transcription start site (Fig. 6) was determined by subjecting the 70-base nuclease-protected fragment to electrophoresis in a high resolution polyacrylamide gel in a lane adjacent to a 'sequencing ladder' of base-specific fragments¹⁹ of the end-labelled *XhoI*-*HinfI* fragment.

We have demonstrated also specific transcription from the *endoH* promoter *in vitro* by run-off transcription experiments in which the purified *XhoI*-*NarI* and *XhoI*-*HinfI* DNA fragments (Fig. 5, top) were used as templates for *in vitro* RNA synthesis. A purified preparation of *ctc*-transcribing RNA polymerase generated transcripts whose lengths (115 and 70 bases for the *XhoI*-*NarI* and *XhoI*-*HinfI* templates, respectively) were in close agreement with those predicted from the position of the 5' terminus of *in vivo* synthesized *endoH* RNA (Fig. 5A). The *veg*-transcribing form of RNA polymerase, in contrast, was inactive with these *endoH*-containing DNAs as templates (data not shown). S₁ nuclease mapping experiments of *in vitro*-generated *endoH* RNA and high resolution gel electrophoresis of the 70 base run-off transcript (data not shown) demonstrated that the start site for transcription *in vitro* was close to or identical with the 5' terminus of *in vivo*-synthesized *endoH* RNA.

We believe that use of the *endoH* promoter is determined by a *ctc*-transcribing form of RNA polymerase, as highly purified preparations of this enzyme used *endoH* efficiently and the elution profiles of *endoH*-transcribing activity and *ctc*-transcribing activity were indistinguishable during gradient elution of RNA polymerase from DNA cellulose (data not shown). The involvement of *ctc*-stimulatory factors in *endoH* transcription, however, has not been established directly by core RNA polymerase complementation experiments with gel-purified proteins. (*B. subtilis* core RNA polymerase cannot use heterologous promoters from bacterial DNA with high G+C base composition; we have not yet devised a protocol for obtaining core enzyme from *S. coelicolor*). There are certain interesting similarities between the *endoH* and *ctc* promoters in positions relative to the transcription start site important in *ctc* use by the Eσ³⁷ form of *B. subtilis* RNA polymerase. Positions in the *ctc* promoter thought to signal *B. subtilis* Eσ³⁷ recognition (judged by homologies to other promoters used by Eσ³⁷ (ref. 10), methylation protection experiments¹⁰ and nucleotide substitution mutations¹¹) are identified in Fig. 6. Although there is little overall homology between the transcription initiation regions of *endoH* and *ctc*, these promoters partially conform to each other (allowing an adjustment of two bp in the spacing between the '-10' and '-35' regions; Fig. 6) at positions functionally important in *ctc* promoter use. These include positions critical for *ctc* use by *B. subtilis* Eσ³⁷ polymerase (Ref. 11; Fig. 6). Thus, *endoH* may be an example of a *ctc*-like promoter in *Streptomyces* whose use is determined by nucleotide sequences similar to those signalling recognition by *B. subtilis* Eσ³⁷ enzyme; this can be investigated by mutational analysis of the nucleotide sequence requirements for *endoH* and *ctc* promoter recognition by *S. coelicolor* RNA polymerase.

Implications

We demonstrate here the existence of at least two distinct forms of RNA polymerase holoenzyme in *S. coelicolor* which differ in

their promoter recognition specificities; these specificities are determined separately by different species of RNA polymerase σ factor. The *S. coelicolor* enzyme containing σ^{35} appeared to be homologous in its specificity to the prototypical eubacterial holoenzyme form, in that it efficiently used a promoter (*veg*) strongly conforming (in its '-10' and '-35' sequences and their spacing) to the canonical structure of the principal class of bacterial promoters²⁰. The enzyme containing σ^{49} (or σ^{37}), on the other hand, was homologous in specificity, judged by its ability to use the *ctc* promoter, to a holoenzyme form ($E\sigma^{37}$) observed so far only in *B. subtilis*, where it is involved in the transcription of developmentally-regulated genes. As evidence for the physiological significance of such holoenzyme forms to the transcription of *Streptomyces* genes, we have shown that highly purified *ctc*-utilizing RNA polymerase (but not purified *veg*-transcribing enzyme) could transcribe from a *Streptomyces* promoter, initiating at a site close to or identical with the start site for *endoH* transcription observed *in vivo*.

RNA polymerase heterogeneity in *S. coelicolor* suggests that the activity of different sets of genes may be determined in part by a variety of holoenzyme forms, each with the capacity to use selectively a characteristic class of promoters. Thus, $E\sigma^{35}$ and $E\sigma^{49}$ may be two members of a larger family of RNA polymerases that determine the differential expression of different gene sets during the course of the *S. coelicolor* life cycle. These gene sets may include but may not be limited to those involved in the differentiation and secondary metabolism of this morphologically complex bacterium. The existence of heterogeneous forms of RNA polymerase is not confined to bacteria with complex life cycles^{21,22}; certain minor holoenzyme forms may determine the expression of gene sets whose activity is unrelated to morphological development, such as heat-shock controlled expression. The cloning of various classes of *Streptomyces* genes, for example drug resistance determinants (see ref. 23), developmental genes (J. M. Piret and K. F. Chater, personal communication) and genes with associated secondary metabolism^{24,25} should furnish other examples of $E\sigma^{35}$ - and $E\sigma^{49}$ -used promoters in this genus and provide templates for identifying additional holoenzyme forms.

Bacteria of the genus *Streptomyces* differ from *Bacillus* species in many obvious morphological and biochemical aspects. *Streptomyces* is filamentous and its DNA is exceptionally high in G+C base composition; it grows into a mycelium and forms spores by septation of aerial hyphae into compartments differentiating into chains of heat-sensitive, resting cells¹. *Bacillus* is unicellular with DNA low in G+C base composition; its spores (endospores) are heat-resistant and formed by a characteristic engulfment process^{3,4}. Nevertheless, as a *B. subtilis* $E\sigma^{37}$ -like form of RNA polymerase exists in *S. coelicolor*, bacteria of both genera may use similar strategies for regulating the expression of developmental genes. Thus, although some of the genes involved in the growth, morphology and differentiation of *Streptomyces* and *Bacillus* must encode very different kinds of structural proteins, the underlying mechanisms of their regulation may be analogous and perhaps even homologous. If this is true, then knowledge of developmental gene regulation strategies in the more extensively studied *Bacillus* genus may elucidate the mechanism of differentiation in *Streptomyces*, only recently begun to be exploited by molecular genetics approaches. For example, the initiation phase of sporulation in *Bacillus* is dependent upon a class of genes known as *spo0* (ref. 26). Mutations in any one of these genes arrest the onset of development and block the production of molecules such as antibiotics and exoenzymes whose synthesis is associated with the earliest stages of differentiation. The *spo0* gene products are thought to act at the level of promoter use by influencing transcription initiation by $E\sigma^{37}$ and other novel forms of *B. subtilis* holoenzyme^{27,28}. The initiation phase of development in *Streptomyces* may be similarly governed by genes in which mutations are highly pleiotropic, preventing the synthesis of certain kinds of secondary metabolites and blocking development at its earliest stage. Indeed, such a class of genes is known—the *bld* genes¹—and

it is conceivable that the products of some of these genes function in similar ways to the *spo0* gene products of *Bacillus*; this could be of practical importance in manipulating the expression of antibiotic pathways in *Streptomyces* species.

Finally, we comment on the remarkable capacity of *Streptomyces* to support the transcription of heterologous genes from a diversity of procaryotic sources. For example, *S. lividans* is an effective host for the use of transcription signals from *E. coli*, *Bacillus* and *Serratia*²⁹, recognizing in at least one instance (*E. coli ampC*) the same '-10' and '-35' signals that promote initiation in *E. coli*³⁰. *Bacillus*, however, is limited in its repertoire of foreign gene expression: for example, *B. subtilis* can use promoters from other *Bacillus* species and from closely-related Gram-positive genera such as *Streptococcus*³¹ and *Staphylococcus*^{32,33}, but promoters from *E. coli*^{7,34,35} or from *Streptomyces* (I. Smith, personal communication) are generally found to be inert in this host, even when they conform to the prototypical sequences for promoter structure. *E. coli* displays an intermediate promoter tolerance pattern, efficiently using many *Bacillus* promoters (at least of the prototypical class) (for example, refs 35, 36), but failing to transcribe from some if not all *Streptomyces* promoters²⁹. These observations suggest that simple conformity to canonical -10 and -35 sequences is not a sufficient condition for signalling transcription in heterologous hosts.

How can we elucidate these patterns of interspecific promoter use, given that all three genera share forms of RNA polymerase capable of recognizing promoters of prototypical structure? Perhaps some general distinguishing feature of *Bacillus*, *E. coli* and *Streptomyces* promoters outside their -10 and -35 regions plays a decisive role in their ability to be used in heterologous hosts. This distinguishing characteristic may be the overall A+T content of the promoter regions of genes from the three genera. For example, *Bacillus* promoters are exceptionally high in A+T content and are sometimes associated with extensive upstream stretches of AT base pairs, important in the efficiency of promoter use^{7,37,38}. *Streptomyces* promoters, at the other extreme, are remarkably deficient in AT base pairs (for example, ref. 39). *E. coli* promoters are intermediate in overall A+T base composition^{20,37}. We suggest, therefore, that the overall high G+C content of *Streptomyces* DNA, including that in and around its promoter sequences, may be irrelevant to promoter use in that genus, but presents a barrier to gene expression in heterologous hosts such as *E. coli* and *Bacillus*, with characteristically lower G+C DNA base composition. In this regard, non-specific barriers to interspecific gene expression, such as DNA base composition, may tend to obscure underlying homologies in the transcriptional machinery of morphologically diverse bacterial genera.

This work was funded by the Eli Lilly Company; we thank Richard Baltz, Paul Burnett and Larry Day for their support. The *endoH* gene, the *ctc* plasmids and the *B. subtilis* RNA polymerases were provided by P. W. Robbins, M. Igo and W. C. Johnson, respectively; we thank them and C. Moran, J. Pero, K. Chater, D. Hopwood and M. Bibb for helpful advice.

Received 20 August; accepted 22 October 1984.

- Chater, K. F. in *Microbial Development* (eds Losick, R. & Shapiro, L.) 89–116 (Cold Spring Harbor Laboratory, New York, 1984).
- Chater, K. & Hopwood, D. A. in *The Biology of the Actinomycetes* (eds Goodfellow, M., Mordarski, M. & Williams, S. T.) 229–286 (Academic, London, 1984).
- Losick, R. & Youngman, P. J. in *Microbial Development* (eds Losick, R. & Shapiro, L.) 63–88 (Cold Spring Harbor Laboratory, New York, 1984).
- Doi, R. H. *A. Rev. Genet.* **11**, 29 (1977).
- Losick, R. & Pero, J. *Cell* **25**, 582–584 (1981).
- Doi, R. H. *Archs Biochem. Biophys.* **214**, 772 (1982).
- Moran, C. P., Jr. et al. *Molec. gen. Genet.* **186**, 339–346 (1982).
- Haldenwang, W. G. & Losick, R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7000–7004 (1980).
- Johnson, W. C., Moran, C. P., Jr. & Losick, R. *Nature* **302**, 800–804 (1983).
- Moran, C. P., Jr., Johnson, W. C. & Losick, R. *J. molec. Biol.* **162**, 709–713 (1982).
- Taiti, K. M. & Moran, C. P., Jr. *J. molec. Biol.* **175**, 285–297 (1984).
- Jones, G. *Archs Biochem. Biophys.* **198**, 195–204 (1979).
- Cooper, C. S. & Chater, K. F. *A. Rev. John Innes Inst.* **66**, 83 (1975).
- Hager, D. & Burgess, R. *Analyt. Biochem.* **109**, 76–86 (1980).
- Robbins, P. W., Wirth, D. F. & Hering, C. *J. biol. Chem.* **256**, 10640–10644 (1981).
- Robbins, P. W. et al. *J. biol. Chem.* **259**, 7577–7583 (1984).
- Tarentino, A. L. & Maley, F. *J. biol. Chem.* **249**, 811–817 (1974).
- Berk, A. J. & Sharp, P. A. *Cell* **12**, 721–732 (1977).
- Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499–559 (1980).
- Rosenberg, M. & Court, D. A. *Rev. Genet.* **13**, 319–353 (1979).

21. Landick, R. et al. *Cell* **38**, 175-182 (1984).
22. Grossman, A. D., Erickson, J. W. & Gross, C. A. *Cell* **38**, 383-390 (1984).
23. Thompson, D. J., Kieser, T., Ward, J. M. & Hopwood, D. A. *Gene* **20**, 51-62 (1982).
24. Chater, K. F. & Bruton, C. J. *Gene* **26**, 67-78 (1983).
25. Malpartida, F. & Hopwood, D. A. *Nature* **309**, 462-464 (1984).
26. Hoch, J. A., Shiflett, M. A., Trowsdale, J. & Chen, S. M. H. in *Spores VII* (eds Chambliss, G. and Vary, J. C.) 131-135 (American Society for Microbiology, Washington, D. C. 1978).
27. Zuber, P. & Losick, R. *Cell* **35**, 275-283 (1983).
28. Gilman, M. & Chamberlain, M. *Cell* **35**, 284-293 (1983).
29. Bibb, M. & Cohen, S. N. *Mol. gen. Genet.* **187**, 265-277 (1982).
30. Jaurin, B. & Cohen, S. N. *Gene* **28**, 83-91 (1984).
31. Yagi, Y., McLellan, T., Frez, W. & Clewell, D. *Antimicrob. Agents Chemother.* **13**, 884-887 (1978).
32. Ehrlich, S. D. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1680-1682 (1977).
33. Kreft, J., Bernhard, K. & Goebel, W. *Mol. gen. Genet.* **162**, 59-67 (1978).
34. Lee, G. & Pero, J. *Mol. gen. Genet.* **180**, 57-65 (1980).
35. McLaughlin, J. R., Chang, S.-Y. & Chang, S. *Nucleic Acids Res.* **10**, 3905-3919 (1982).
36. Gray, O. & Chang, S. J. *Bacteriol.* **145**, 422-428 (1981).
37. Lang, N. thesis, Harvard Univ. (1982).
38. Banner, C. D. B., Moran, C. P., Jr. & Losick, R. *J. molec. Biol.* **168**, 351-365 (1983).
39. Thompson, C. J. & Gray, G. S. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5190-5194 (1983).
40. Shorestein, R. G. & Losick, R. *J. biol. Chem.* **248**, 6163-6169 (1973).
41. Penn, M. D., Thireos, G. & Greer, H. *Molec. cell. Biol.* **4**, 520-528 (1984).

Molecular genetics of *Krüppel*, a gene required for segmentation of the *Drosophila* embryo

Anette Preiss, Urs B. Rosenberg, Andrea Kienlin, Eveline Seifert & Herbert Jäckle

Max-Planck-Institut für Entwicklungsbiologie, Spemannstr. 35, 7400 Tübingen, FRG

Krüppel is a member of the 'gap' class of segmentation genes of *Drosophila melanogaster*, mutations of which cause contiguous groups of segments of the fruitfly embryo to fail to develop. In the case of *Krüppel* mutant embryos, thoracic and anterior abdominal segments are deleted. The molecular cloning of the *Krüppel* locus will lead to an understanding of the crucial role that gap genes seem to have in early embryonic development. It has already allowed the identification of a blastoderm-specific *Krüppel* transcript and the phenotypic rescue of mutant embryos by injected cloned DNA.

SEGMENTATION is one of the most striking features of the *Drosophila* larva. The segmental pattern is established at the blastoderm stage and requires the activity of both maternally and zygotically active genes¹⁻⁵. Segment primordia become visible at the end of gastrulation, their number corresponding to the number of segments later seen in the larva, where segmentation and the identity of segments are defined by the pattern of cuticular markers (Fig. 1a). Recent systematic searches for zygotically active genes led to the identification of about 20 loci affecting embryonic segmentation⁶⁻⁹. In mutant embryos of the three 'gap' loci, continuous stretches of adjacent segments are deleted⁹. The gap in the pattern occurs in specific morphologically-defined subregions of the larva. In *hunchback*, the thoracic segments are deleted, whereas in *knirps* only two rather than eight denticle bands are formed in the abdomen⁹. The largest gap in the segment pattern is observed in homozygous *Krüppel* (*Kr*) embryos⁹⁻¹¹. Strong *Kr* alleles lack all thoracic and anterior abdominal segments, which are replaced by a mirror-image duplication of the normal posterior abdominal segments (Table 1, Fig. 1b). In intermediate *Kr* alleles, the thorax and four abdominal segments are deleted (Table 1, Fig. 1c). Weak alleles develop a normal head, prothorax and up to seven abdominal segments, leaving small but variable gaps (Table 1, Fig. 1d). Neither intermediate nor weak alleles show any sign of mirror-image duplications.

The requirement for *Kr*⁺ gene activity is strictly zygotic and appears to be confined to early embryogenesis¹¹. Maternal dosage of *Kr*⁺ has no effect on the embryonic phenotype, nor does homozygosity for *Kr* prevent germ cells from making normal eggs capable of normal embryonic development when fertilized by wild-type sperm¹¹. Homozygous *Kr* embryos can be distinguished from normal wild-type or heterozygous *Kr* embryos shortly after the beginning of gastrulation. Therefore, the *Kr*⁺ gene must be transcribed after egg deposition at stages earlier than gastrulation.

Cytological localization

Two deficiencies (*B80* and *Kr*¹, Table 1) placed the *Kr* locus into polytene chromosomal region 60F2-5 at the tip of the right arm of the second chromosome, very close to the dominant gain of function mutant *Irregular facets* (*If*; map position 2-107.6)¹¹.

Reversion by X-ray mutagenesis of the *If* mutation to normal eye morphology located chromosomal rearrangements in the *Kr* region using a simple F₁ screen. In about 20,000 chromosomes tested, we found four *If* revertants, none of which complemented *Kr*. We isolated also 12 new ethyl methane sulphonate (EMS)-induced *Kr* alleles by screening for mutations which did not complement the embryonic phenotype. We therefore had a total of 26 available *Kr* alleles, 5 of which were X-ray-induced, 19 EMS-induced and two spontaneously arising. Seven of these mutations (five deficiencies, one translocation (*AK1*) and one inversion (*UR1*)) were associated with visible chromosomal rearrangements (Table 1). Cytological analysis of these rearrangements allowed us to map *Kr* to the polytene chromosome band 60F3. One of the *If* revertants (*API*), which by genetic criteria behaved as a *Kr* deletion, showed no detectable cytological abnormality. This chromosome was found to contain a small deletion (see below) that proved useful in the analysis of cloned DNA segments.

Microcloning of chromosome band 60F3

In order to clone the *Kr* gene, we microdissected a single polytene chromosome band (60F3) from salivary gland chromosome squashes obtained from the Oregon R strain of *Drosophila*. DNA was extracted from a total of four bands and digested to completion with *EcoRI* for ligation with λ 641 vector DNA. The resultant DNA was packaged *in vitro* and recombinant phage particles were selected on the *Escherichia coli* pop 13^b strain¹². From a total of 175 recombinant phage clones, 122 were identified which had low or single-copy number DNA inserts, as judged by plaque hybridization analysis using nick-translated, ³²P-labelled *Drosophila* wild-type DNA as a probe for repetitive sequences.

DNA from 72 clones was analysed by *EcoRI* restriction analysis on agarose gels. About 80% of the clones had inserts in the range 0.5-8 kilobases (kb), the average length being 2-3 kb. To identify clones containing DNA sequences that were absent in the *API* deficiency, inserts with single-copy DNA were purified, ³²P-labelled by nick-translation and hybridized to either dot blots or Southern blots containing *EcoRI*-digested genomic DNA from embryos homozygous for *B80*, *Kr*¹ or *API*, or with DNA from wild-type Oregon P2 embryos. The first DNA

Fig. 1 Cuticular pattern of *Drosophila* larvae. *a*, Dark-field photograph of a wild-type larva; *b*, strong phenotype *Kr*¹ embryo; *c*, intermediate phenotype *Kr*^{IV} embryo; *d*, weak *Kr* phenotype *Kr*^{XII} embryo. T1, pro-; T2, meso-; T3, metathorax. A 1-8, abdominal segments. Note the duplication of the sixth abdominal segment with reversed polarity in the strong *Kr* phenotype. *e*, Phase-contrast photograph of the enlarged abdominal region of a *Kr*¹ homozygous embryo injected with clone ER3 DNA (see text and Table 2). Note the phenotypic rescue of this embryo, indicated by the number of segments adjacent to the sixth abdominal segment (A¹⁻³). Note also the normal polarity (indicated by arrows) of the additional segments never seen in the strong *Kr* phenotype (see *b*). Magnifications in *a* and *b-e* are different (as can be judged from the size of denticle bands).

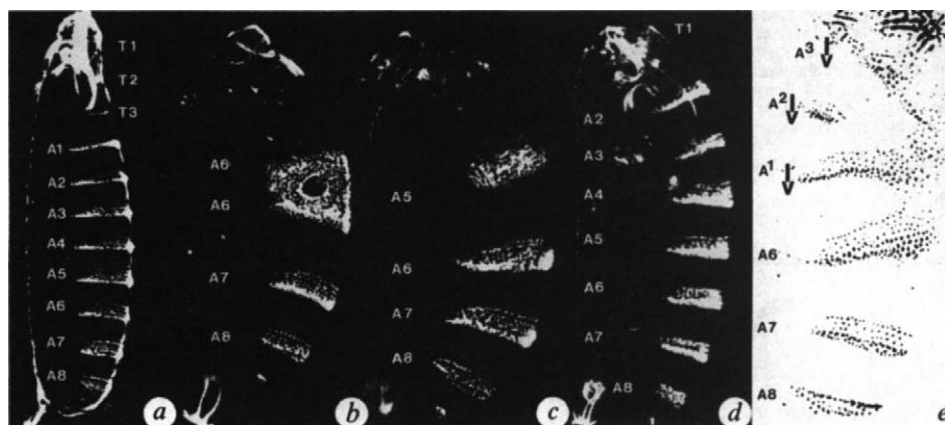


Table 1 Analysis of *Kr* mutations

Allele	Parental chromosome	Mutagen	Morphology of <i>Kr</i> phenotype	Cytology	Molecular analysis	Ref.
<i>Kr</i> ¹	U	S	ST	Df 60 F2-5	Deletion	9-11
<i>Kr</i> ²	C	E	ST		Normal	11
<i>Kr</i> ³	C	E	ST		Normal	11
<i>B80</i>	U	X	ST	Df 60 F2-5	Deletion	11, 20
<i>Kr</i> ^{6A69}	U	E	ST	Df 60 F3-5	Deletion	11
<i>Kr</i> ⁷	<i>If</i>	E	ST		Normal	8
<i>Kr</i> ⁸	<i>If</i>	E	ST		Normal	This work
<i>Kr</i> ⁹	<i>If</i>	E	ST		Normal	This work
<i>SB1</i>	<i>B1 If</i>	X	ST	Df 60E10-F5*	Deletion	This work
<i>UR1</i>	<i>B1 If</i>	X	ST	In 58A-60F3	Deletion	This work
<i>API</i>	<i>B1 If</i>	X	ST	Normal	Deletion	This work
<i>AK1</i>	<i>If</i>	X	ST	T(Y;2)60 F3-5†	Translocation	This work
<i>Kr</i> ¹¹	M	S	ST		Deletion	This work
<i>Kr</i> ^{L14}	U	E	ST	Df 60 F2-5	Deletion	8
<i>Kr</i> ^{III} ^{A102}	C	E	I		Normal	7
<i>Kr</i> ^{IV}	<i>If</i>	E	I	DS: T 1-3; A 1-4	Normal	This work
<i>Kr</i> ^{VI}	<i>If</i>	E	I	Dup: not observed	Normal	This work
<i>Kr</i> ^I ⁰⁶⁶	C	E	W	DS: T 2, 3; A 1, 2	Normal	7
<i>Kr</i> ^{II} ^{ES7}	C	E	W	DS: T 2, 3; A 1, 2	Normal	7
<i>Kr</i> ^V	<i>If</i>	E	W	DS: T 2, 3; A 1-4; V: A 4	Normal	This work
<i>Kr</i> ^{VII}	<i>If</i>	E	W	DS: T 2, 3; A 1-3; V: A 3	Normal	This work
<i>Kr</i> ^{VIII}	<i>If</i>	E	W	DS: T 2, 3; A 1-3; V: A 3	Normal	This work
<i>Kr</i> ^{IX}	<i>If</i>	E	W	DS: T 2, 3; A 1, 2; V: A 2, 3	Normal	This work
<i>Kr</i> ^X	<i>If</i>	E	W	DS: T 2, 3; A 1-3	Normal	This work
<i>Kr</i> ^{XI}	<i>pr</i>	E	W	DS: T 2, 3; A 1-3	Normal	This work
<i>Kr</i> ^{XII}	<i>If</i>	E	W	DS: T 2, 3; A 1; V: A 2	Normal	This work

Seventeen new *Kr* mutants were isolated in screens over the original *Kr*¹ allele of Gloor¹⁰ using EMS mutagenesis⁶⁻⁸ or X-ray (4,000R) mutagenesis as described in the text. Parental chromosomes: U, unknown origin; C, *cn bw sp*; *If*, *b pr cn wxt If*; *B1 If*, *b pr B1 c If*; M, M(2)H⁵⁵; *pr*, *b pr If*. Note that the chromosomes designated *If* and *pr* were obtained by recombination of *B1 If*, so that all three chromosomes should be identical in the *If* region. A detailed description of the genetic markers used in this study as well as the balancer chromosomes *CyO* and *SM1* used for stock-keeping are described in ref. 20. Mutagens: S, spontaneous origin; E, EMS-induced; X, X-ray-induced. Phenotypes of homozygous or trans-heterozygous *Kr* embryos were scored in cuticular preparations (for methods, see refs 6-9, 11). ST, Strong *Kr* phenotype (see Fig. 1b); I, intermediate phenotype (see Fig. 1c); W, weak phenotype (see Fig. 1d). T 1-3, thoracic segments; A 1-8, abdominal segments; rev, reversed polarity in a given segment; DS, deleted segment(s); Dup, duplicated segment(s); V, variable pattern element observed in low frequency. Cytology: polytene chromosomes were prepared from salivary glands, stained with orcein and analysed at $\times 400$ magnification. Df, deficient; *SB1* is deficient for *gsb* (ref. 7) and *Kr* but not for *zip* (ref. 7). *AK1* was found as a mosaic that had only one normal eye, but stabilized to the *If* revertant phenotype in the next generation. Note that reversion of *If* is never complete and is strongly dependent on the chromosomal background. The four *If* revertants (*UR1*, *SB1*, *API*, *AK1*) were originally recovered over *SM1* chromosomes; they had completely normal eyes. In *trans* over *CyO* or any chromosome of Oregon R origin, however, eyes are still rough and contain some irregular facets (A.P. *et al.*, unpublished data). *Kr*¹¹ was a gift of C. Nüsslein-Volhard.

* The deficient chromosome segment is replaced by translocated chromosome bands probably derived from chromosome 4.

† Y-chromosome translocation involving the region distal to 60 F2.

segment of interest was found in clone 57; it hybridized either to a 9-kb fragment of wild-type DNA or to a smaller 3.6-kb fragment of *API* DNA, but not to *Kr*¹ and *B80* DNA (Fig. 2a,b). *Kr*¹ and *B80* DNA are deficient in sequences of clone 57; clone 57 maps at or near to a deficiency breakpoint of the *API* chromosome. These findings were confirmed by Southern

blot analysis of homozygous *API* DNA, heterozygous *API* DNA and the DNA from flies homozygous for *If*, the chromosome used for mutagenesis (Fig. 2c).

Clone 57 was used to initiate a chromosomal 'walk'¹³ by screening a Canton S library¹⁴. We also screened libraries prepared from DNA of Oregon R embryos and from flies homozy-

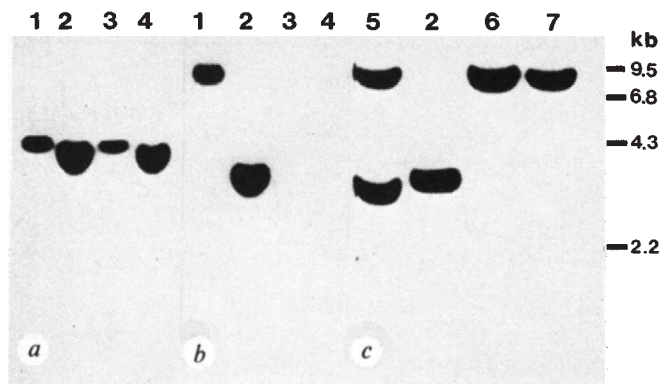


Fig. 2 Localization of microcloned sequences within cytological *Kr* deficiencies. *a*, Southern blots of *Eco*RI-digested genomic DNA. The loading of DNA in the different lanes was checked by probing with a single-copy DNA clone from outside the *Kr* region. *b*, Rehybridization of the same filter (after removal of old hybridization signals) using the clone 57 insert as a probe. Note the absence of bands in lanes 4 and 5 and the different sizes of bands in lanes 1 and 2. *c*, Southern blots containing similar amounts of *Eco*RI-digested DNA from homozygous *API* embryos, heterozygous *API* flies and from flies containing the parental chromosomes used for mutagenesis.

Methods: DNA was used from wild-type embryos (*a*, *b*, lanes 1), embryos homozygous for *API*, (*a*, *b*, *c*, lanes 2), *Kr*¹ (*a*, *b*, lanes 3), *B80* (*a*, *b*, lanes 4) and from flies *API*/*SM1* (*c*, lane 5), *If*/*If* (*c*, lane 6) and *B1 If*/*SM1* (*c*, lane 7) (ref. 21; see Table 1 for details). Homozygous *Kr* embryos were individually collected under the stereomicroscope. To avoid losses of DNA, RNA was not removed from the DNA preparations. Consequently, the amount of DNA loaded onto the gel was checked as described for *a*. Restriction digests, gel preparation, Southern blotting, hybridization and nick-translation of probes were performed as described previously^{13,14,21}.

gous for *If*. For the initial step, insert DNA from the phages was cross-hybridized to detect the overlapping sequences homologous to the clone 57 insert. Restriction fragments from either side of the region of overlap were purified and used to continue the walk in both directions.

Molecular mapping

We constructed a physical map of over 50 kb of the *Kr* region, using overlapping DNA clones obtained by chromosomal walking. The *Kr* region is contained in 39 kb of single-copy DNA surrounded by blocks of repetitive DNA (Fig. 3). On one side, the region is flanked by ~12 kb of tandemly repeated 0.8-kb *Eco*RI subunits (+34 to +46), and on the other by repetitive elements present at a low frequency in the genome, and showing heterogeneity in wild-type strains (-5 to -15; A.P., unpublished data). We found a low level of restriction site polymorphism in different strains (Fig. 3). Inserted elements not affecting *Kr*⁺ gene function were observed at three different sites. One insertion, at position +32, was present in all the Canton S clones analysed but was absent from both *If* and Oregon R clones. The two additional insertions at position +1 and +28 were found in 1 out of 20 and 10 clones, respectively (data not shown), indicating a low degree of polymorphism in the DNA used for library construction.

Twenty-six *Kr* mutations (see Table 1) were analysed by hybridizing ³²P-labelled restriction fragments of region -5 to +12.7 to Southern blots of genomic DNA from homozygous embryos and from flies possessing the *Kr* allele in *trans* to CyO balancer chromosomes. To analyse region 0 to +9.5 (see examples in Fig. 4a) in heterozygous DNA, we used DNA from the CyO balancer chromosomes, which lacked two restriction sites (*Sal*I, *Hind*III) present in DNA of *Kr* alleles originating from *If* (see Fig. 3). In ambiguous cases, however, we tested DNA prepared from individually scored homozygous *Kr* embryos, to confirm or extend the results of the above analysis

(for example, see Fig. 2). Of the 14 strong alleles analysed for region -12 to +33, nine were associated with DNA deletions (Fig. 3) and were grouped into two classes. One class consists of deletions for the entire cloned single-copy DNA region from -5 to +34. This class includes *B80*, *Kr*¹, *Kr*^{L14} and *SBI*, classified previously as visible cytological deficiencies (Table 1). The entire region was also deficient in *AK1* females, as the DNA was translocated onto the Y chromosome (Table 1, Fig. 3). The second class of deletions contained breakpoints within the cloned region. The breakpoints of the three deletions *API*, *Kr*^{6A69} and *Kr*^{JI} were found close to positions +1, -5 and +5, respectively, and DNA sequences extending up to map position +34 (or beyond) were absent from all three mutants. The smallest deletion uncovered the region from ~0 to +9 and was associated with one of the inversion breakpoints of chromosome *UR1*. In summary, all of the X-ray-induced *Kr* mutations, both of the spontaneous mutants and two of the EMS-induced *Kr* alleles have associated deletions of DNA (see Table 1). No evidence for DNA deletions or other DNA rearrangements was found in the other *Kr* mutations.

Trans-heterozygous *Kr*^{JI}/*UR1* embryos show the strong *Kr* phenotype. On the DNA level, the region of overlap of the two deficiencies is less than 4 kb of DNA. This indicates that DNA sequences from map positions +5 to +9 contain parts of the *Kr*⁺ gene or, as a minimum, sequences essential for normal *Kr*⁺ gene expression.

Kr transcription in the blastoderm

Homozygous *Kr* embryos can be distinguished from heterozygous and wild-type embryos shortly after the beginning of gastrulation¹¹. Because the *Kr* phenotype has no maternal component, the *Kr*⁺ gene must be active after egg deposition. The molecular data indicate that *Kr* should be transcribed close to or from the region characterized by the 4-kb DNA deletions (Fig. 3). We therefore used cloned DNA segments corresponding to region -5 to +17 to probe Northern blots loaded with equal amounts of poly(A)⁺ RNA from various stages of *Drosophila* development. Region 0 to +9.5 hybridizes to a 2.5-kb poly(A)⁺ RNA present only in 2-5 h-old embryos (Fig. 4b). Titration experiments using actin messenger RNA as a standard show that this transcript amounts to ~0.01% of total poly(A)⁺ RNA. We could no longer detect the transcript in extended germ-band-stage embryos, even though our Northern blots are capable of detecting transcripts at a 10-fold lower concentration. No poly(A)⁻ transcripts were detected. (The detailed study of *Kr*⁺ transcription will be described elsewhere; U.B.R. *et al.*, in preparation.)

Both molecular analysis and transcript mapping are consistent with the localization of the *Kr* locus in or close to region 0 to +10. To identify *Kr*⁺ gene activity, we injected cloned DNA into mutant embryos and scored them for possible 'rescue' effects resulting from *Kr*⁺ gene activity provided by the injected DNA.

Phenotypic rescue of *Kr* mutants

The *Kr* alleles can be ordered into an allelic series where the size of the gap in segmentation is proportional to the strength of the mutant alleles. In amorphic alleles, all three thoracic and five anterior abdominal segments are deleted and replaced by a mirror-image duplication of the normal posterior abdomen. Weaker alleles have no mirror-image duplication and progressively fewer segments are deleted (Fig. 1b-d, Table 1), indicative of residual *Kr*⁺ gene activity in the different *Kr* mutations¹¹. The injection of cloned DNA provided low *Kr*⁺ activity which partially normalizes the amorphic phenotype.

DNAs from various clones for most of the cloned region were injected into embryos of *Ddc Kr*¹/*SM1* parents. The dopa decarboxylase mutation (*Ddc*) allowed *Kr*¹/*Kr*¹ homozygous embryos to be unequivocally distinguished from their siblings, as it renders the cuticle and mouth parts of homozygous *Kr*¹ larvae unpigmented^{11,15}. After injection of DNA from clone ER3, the segment anterior to the sixth abdominal segment of homozygous *Kr*¹ embryos often showed normal polarity (Table

Fig. 3 Molecular map of the *Kr* region. The coordinates (in kb) are based on an *EcoRI* site at the start point of the chromosomal walk with clone 57 (see Fig. 2). The small size of this clone is probably due to cutting at the *EcoRI** site under the high glycerol concentrations used for microcloning¹². **a**, Restriction map of DNA from the *If* chromosome used for mutagenesis (Table 1). The same map was found for the *B1 If* chromosomes used to make chromosomal rearrangements (Table 1). The four restriction enzymes used for mapping were: R, *EcoRI*; B, *BamHI*; S, *SalI*; H, *HindIII*. **b**, Individual phage clones covering the *Kr* region. The cloned DNA originated from homozygous *If* flies (designation AI, each insert ends with a true *EcoRI* site), Canton S embryos¹⁴ (designation cc), or Oregon R embryos (designation ER; each insert ends with a *Sau3A* site due to partial digestion of genomic embryo DNA and ligation into the *BamHI* site of vector DNA). Note polymorphic restriction sites in Oregon R DNA, that is, two additional *EcoRI* sites and the absence of one *SalI* and one *HindIII* site also absent in *CyO* DNA (see text). Fragments smaller than 0.5 kb and other clones overlapping the indicated clones are not shown. Vectors used for library construction were EMBL 4 (ER), Charon 4 (cc) and Charon 4A (AI). **c**, Deletions in *Kr* mutations. Of the 26 *Kr* alleles analysed (see Table 1), 9 represent deletions larger than 500 base pairs (bp). *AK1* (dashed line), translocation of the entire cloned region to the Y chromosome (data not shown). Solid bars, deficient DNA sequences; open bars, limits of the break point mapping. Note the small 4-kb deletion of *Kr*^{J1}/*UR1* embryos which show the strong *Kr* phenotype.

Methods: DNAs from phages were mapped with four restriction enzymes after separation on 0.8% agarose gel. End fragments of DNA inserts were labelled by nick-translation to screen the three libraries for homologous sequences flanking either side of the original clone. For each walking step, the presence of repetitive sequences was tested by reverse Southern analysis using ³²P-labelled genomic DNA as a probe²¹. As an additional control, the size of each restriction fragment was checked using Southern blots of genomic DNA.

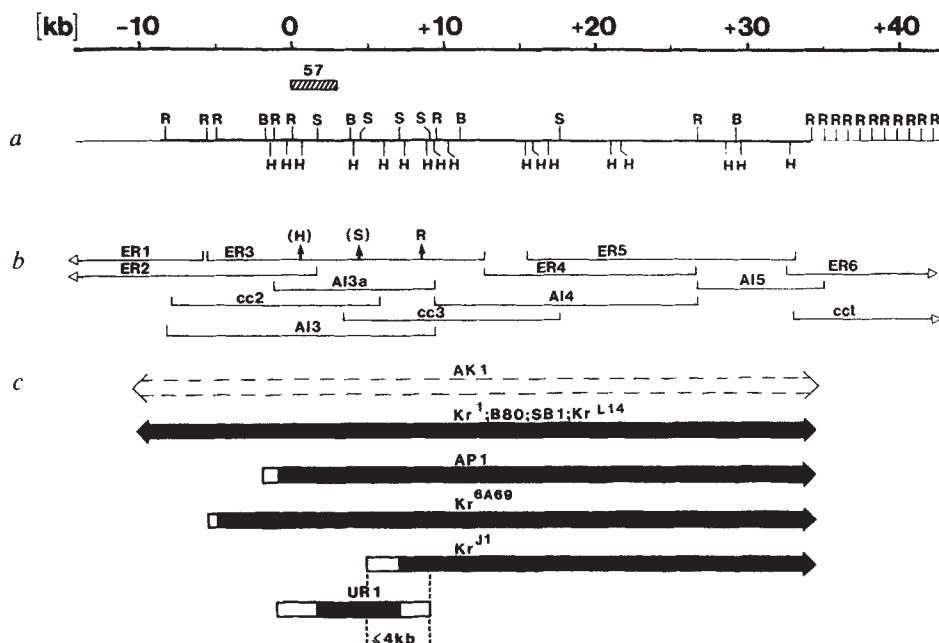


Table 2 Phenotypic rescue of mutant *Kr* embryos injected with cloned DNA

Origin of injected DNA (map position)	Site of injection (% e.l.)	Concentration of injected DNA ($\mu\text{g ml}^{-1}$) (>10 experiments)	No. of homozygous <i>Kr</i> embryos injected	No. of injected homozygous <i>Kr</i> embryos which develop (%)			
				Strong phenotype	Intermediate phenotype	Weak phenotype	Involted (normal) head
cc2 (-7.9 to +5.7)	50-60	130	111	111 (100)	0	0	0
AI3a (-1.2 to +9.5)	50-60	130	162	162 (100)	0	0	0
AI3 (-6.8 to +9.5)	50-60	130	66	66 (100)	0	0	0
ER3 (-5.1 to +12.7)	50-60	130	115	63 (55)	46 (40)	6 (5)	3 (3)
cc3 (+3.3 to +17.6)	50-60	130	128	128 (100)	0	0	0
ER5 (+15.3 to +33.3)	50-60	130	102	102 (100)	0	0	0
ER3 (-5.1 to +12.7)	50-60	50	70	67 (96)	3 (4)	0	0
	50-60	80	78	60 (77)	18 (23)	0	0
	50-60	150	59	37 (63)	22 (37)	2 (3)	0
	50-60	200	50	48 (96)	2 (4)	0	0
	50-60	400	60	60 (100)	0	0	0
ER3 (-5.1 to +12.7)	0-40	130	71	71 (100)	0	0	0
	60-80	130	90	79 (88)	11 (12)	0	0
	80-100	130	59	58 (98)	1 (2)	0	0

DNA was extracted from phages using standard methods. Their position in the *Kr* region corresponds to map positions in Fig. 3. Embryos were injected in lateral positions along the anterior-posterior axis of the embryo (% e.l., % egg length; posterior is 0%). Homozygous *Kr* embryos (25% of injected embryos) show the *Ddc* phenotype and are therefore distinguishable from their heterozygous and wild-type siblings. Normal embryos or embryos heterozygous for *Kr* were scored as pigmented embryos¹¹. Embryos with defects other than segment fusion or small cuticular holes were scored as undifferentiated (14-30% of injected embryos). The strong *Kr* phenotype is indistinguishable from that of uninjected *Kr* embryos; 'intermediate' phenotype refers to injected embryos that develop a phenotype similar to the phenotype of *Kr* alleles III, IV or VI, and 'weak' phenotype refers to injected embryos that develop more than four abdominal segments with normal polarity (for example, see Fig. 1e). For the classification of phenotypes see Table 1 and Fig. 1.

Methods: Eggs from a *Ddc*ⁿ⁷*Kr*¹/*SM1* mating were collected on apple juice agar plates. After dechoriation, embryos were mounted on coverslips, dried on silica gel and covered with oil. DNA was dissolved in injection buffer and injected with a 5 μm -wide pipette¹⁶. Embryos were injected at about pole cell stage, as this was found to be most efficient for rescue response. After injection, embryos were allowed to develop for 2 days in a moist chamber at 20°C. Normally, 70-90% of the injected normal embryos hatched. After embedding, the cuticular pattern of unhatched embryos was analysed under the phase-contrast microscope¹¹.

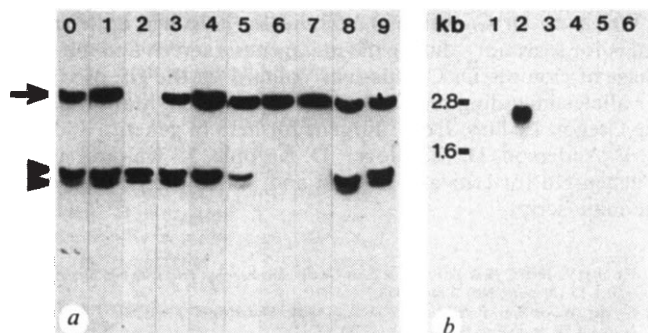


Fig. 4 *a*, Mapping of chromosomal deficiencies in the 0–9.5 map region using DNA of heterozygous *Kr/CyO* flies. Due to three polymorphic restriction sites (see Fig. 3), differences in DNA of *CyO* and *Kr* chromosomes can be distinguished in heterozygous DNA. For example, Southern blots containing *SalI*-digested genomic DNA from heterozygous flies were probed with a nick-translated, ^{32}P -radiolabelled 2.3-kb subclone (map position +3.1 to +5.4). This clone hybridizes to a 5.4-kb *SalI* fragment (arrow) in *CyO* DNA of balancer chromosomes and two (2.6 and 2.8 kb) fragments (arrowheads) in *If* chromosomes. Note the absence of the lower bands in lanes 6 and 7, indicating the lack of these sequences in *URI* and *API* DNA, respectively. The absence of the 2.6-kb fragment in lane 4 (*Kr¹*, overexposed bands) indicates the break point within this fragment close to map position +4.3. The weak intensity of the lower bands in *AKI* DNA (lane 5) is consistent with genetic and cytological data indicating a Y-translocation of this region (Table 1). As the translocated region is present only in males (about half the flies), the corresponding bands amount to one-quarter of the copy number in fly DNA. The *CyO* fragment was absent in *If* homozygous DNA (lane 3), but present in any combination with *CyO*. For methods, see legends to Figs 2 and 3. Lane 0, *Kr^{VI}*; lane 1, *Kr⁹*; lane 2, parental *If/If*; lane 3, parental *If/CyO*; lane 4, *Kr¹*; lane 5, *AKI*; lane 6, *URI*; lane 7, *API*; lane 8, *Kr²*; lane 9, *Kr³*. For detailed description of the phenotypes see Table 1. *b*, Developmental profile of the *Kr⁺* gene transcript. Northern blots of poly(A)⁺ RNA from different stages of *Drosophila* development were probed with cloned DNA corresponding to region 0 to +9 on our map (Fig. 3). Lane 1, early cleavage stage (0–60 min); lane 2, blastoderm stage and gastrulation (2–5 h); lane 3, late gastrulation to hatching (8–24 h); lane 4, third-instar larvae; lane 5, pupae; lane 6, adults.

Methods: Staged embryos were collected from population cages. RNA was prepared using the guanidinium isothiocyanate method²². Poly(A)⁺ RNA was separated on oligo(dT) columns as described previously²³. From each stage, 10 μg of denatured poly(A)⁺ RNA was separated on 1.2% agarose gels together with RNA and DNA size markers. Northern blots and hybridization were as described previously²⁴.

2); these embryos more closely resembled the intermediate rather than the strong *Kr* phenotype of uninjected *Kr¹* embryos. *Kr¹* embryos injected with DNA from other clones of the region showed only the extreme phenotype (Table 2). In addition to the 40% of intermediate phenotype *Kr¹* embryos found (Table 2), a small fraction of *Kr¹* embryos (5%) developed more than two additional segments with normal polarity (Fig. 1e), indicative of a substantial weakening of the amorphic phenotype.

The phenotypic rescue response of *Kr¹* embryos was observed only when DNA from the ER3 clone was injected into a middle position of the embryo (Table 2). Furthermore, the rescue response was dependent on both injection earlier than blastoderm formation and concentration of injected DNA. The response was linear below 130 $\mu\text{g ml}^{-1}$ DNA, but decreased at higher concentrations (Table 2). The injection experiments map *Kr⁺* gene activity and, therefore, the entire *Kr* transcription unit to the ER3 clone which overlaps region –5.1 to +12.7 on our map. This clone includes both the DNA with the 4-kb deletion that had been genetically mapped as an essential region of the *Kr⁺* gene, and the DNA from which the 2.5-kb poly(A)⁺ RNA present in 2–5 h-old embryos is transcribed. Smaller clones, AI3 and AI3a (see Fig. 3; Table 2), did not evoke phenotypic rescue,

indicating that DNA sequences at map position +9.5 to +12.7 contain functional parts of the *Kr⁺* gene.

Discussion

We report here the molecular cloning of *Kr*, a segmentation gene of *Drosophila*. Clones obtained from microdissected DNA facilitate the isolation of some 50 kb of genomic DNA in a series of overlapping clones, enabling us to identify DNA sequences within a 4-kb interval as those required for *Kr⁺* gene function. One clone containing this 4-kb DNA segment and 14 kb of flanking sequences contained *Kr⁺* activity, as judged by phenotypic rescue following injection of the cloned DNA into mutant embryos. The functional *Kr* locus was therefore mapped to a DNA sequence of maximum length 18 kb. The DNA sequences corresponding to cloned DNA providing the biological *Kr⁺* gene activity, code for a 2.5-kb poly(A)⁺ RNA first detected at blastoderm stage.

Phenotypic rescue of the mutant *Kr* embryos after injection of cloned DNA allowed the *Kr* sequence to be localized and positively identified within the cloned region. This direct biological assay was suggested by earlier experiments¹⁶ showing transcription of injected cloned DNA in developing embryos. About 80% of the DNA was degraded in the cytoplasm; transcription of the remaining DNA was more than one order of magnitude below the efficiency of transcription from endogenous genes. In our most successful rescue experiments, we injected some 300 pl of DNA at a concentration of 130 $\mu\text{g ml}^{-1}$ or $\sim 10^6$ molecules per injected embryo (100 times the number of *Kr⁺* gene copies of normal blastoderm stage embryos). We expected transcription of the injected genes to be less efficient than that of endogenous genes in heterozygous embryos and the resulting biological activity to be even lower. As viable heterozygous embryos show a distinct thoracic abnormality¹¹, the degree of rescue response obtained was in the expected range; thus, *Kr⁺* activity was still below 50% of the normal activity. An increase in DNA concentrations above 150 $\mu\text{g ml}^{-1}$ led to a decrease in rescuing activity. We assume that highly concentrated concatenated DNA simply aggregates in the embryo and thus prevents a substantial portion of the DNA from entering the proper cellular compartment. We have frequently observed translucent regions in developing embryos at the site of injection, which may correspond to highly concentrated injected DNA.

The observed phenotypic rescue of mutant embryos suggests that the ER3 clone contains all the regulatory sequences required for spatial and temporal regulation of endogenous *Kr⁺* gene expression. In wild-type embryos, *Kr* is transcribed at the blastoderm stage, when the rate of embryonic transcription is increased dramatically¹⁷. Random transient ER3 DNA expression may produce a low level of transcripts during the period of development, which normally requires *Kr⁺* gene activity. Sequences important for endogenous *Kr⁺* activity could thus be outside the ER3 fragment; this would provide an alternative explanation for the weak rescue response. Other overlapping clones (see Table 2) (including the transcribed region) failed to give a rescue response. As two of these clones, AI3 and AI3a, contain all the ER3 sequence except for the 3.2-kb segment at map position +9.5 to +12.7, essential sequences for *Kr⁺* activity must lie within this 3.2 kb. If a phenotypic indicator of rescue is available (as in *Kr* embryos), this assay for the function of cloned DNA is much easier than genetic transformation experiments which require transposon-mediated introduction of cloned DNA into the *Drosophila* germ line¹⁸ and genetic analysis of the transformants.

The rescue of the *Kr* phenotype required DNA injection in the middle of the embryo. This region of the embryo coincides with the anlage of thorax and anterior abdomen on the blastoderm fate map¹⁹, and appears to be the most sensitive to the absence of the *Kr⁺* gene product, as reflected in the lack of thoracic and anterior abdominal segments in the weak *Kr* alleles. The results of the injection experiments may be the first indication of localized *Kr⁺* gene product requirement during the

blastoderm stage. The early period when *Kr* is transcribed coincides with the time of cell commitment long before segments become visible. This suggests that the *Kr*⁺ function is necessary for the initiation rather than the maintenance of thoracic and abdominal segmentation in the early *Drosophila* embryo, in agreement with the earlier conclusion that the absence of *Kr*⁺ activity can result in a change in the fate of primordial cells in *Kr* embryos¹¹.

Received 20 August; accepted 2 October 1984.

1. Lawrence, P. A. *J. Embryol. exp. Morph.* **30**, 681-699 (1973).
2. Wieschaus, E. & Gehring, W. J. *Dev. Biol.* **50**, 249-263 (1976).
3. Schubiger, G. *Dev. Biol.* **50**, 476-488 (1976).
4. Capdevila, M. P. & Garcia-Bellido, A. *Wilhelm Roux Arch. dev. Biol.* **179**, 125-143 (1976).
5. Szabad, J., Schüpbach, T. & Wieschaus, E. *Dev. Biol.* **73**, 265-271 (1979).
6. Wieschaus, E., Nüsslein-Volhard, C. & Jürgens, G. *Wilhelm Roux Arch. dev. Biol.* **193**, 296-307 (1984).
7. Nüsslein-Volhard, C., Wieschaus, E. & Kluding, H. *Wilhelm Roux Arch. dev. Biol.* **193**, 267-282 (1984).
8. Jürgens, G., Kluding, H., Nüsslein-Volhard, C. & Wieschaus, E. *Wilhelm Roux Arch. dev. Biol.* **193**, 283-295 (1984).
9. Nüsslein-Volhard, C. & Wieschaus, E. *Nature* **287**, 795-801 (1980).
10. Gloor, H. *Arch. Julius Klaus-Stift.* **25**, 38-44 (1954).
11. Wieschaus, E., Nüsslein-Volhard, C. & Kluding, H. *Dev. Biol.* **104**, 172-186 (1984).

We thank Dr C. Nüsslein-Volhard for help and criticism, S. Baars for assistance during the mutagenesis screen and the early phase of cloning, Dr C. Nüsslein-Volhard for the gift of several *Kr* alleles including the *Ddc* marked *Kr*¹, Dr A. Mahowald for the Oregon P2 flies, Dr G. Jürgens for help in genetics and Drs K. V. Anderson, D. M. Glover, D. Knipple, K. Kalthoff and J. Wachenfeld for critically reading and Ms C. Blume for typing the manuscript.

12. Pirrotta, V., Jäckle, H. & Edström, J. E. in *Genetic Engineering: Principles and Methods* Vol. 5, 1-17 (Plenum, New York, 1983).
13. Bender, W., Spierer, P. & Hogness, D. S. *J. molec. Biol.* **168**, 17-33 (1983).
14. Maniatis, T. et al. *Cell* **15**, 687-701 (1978).
15. Wright, T. R. F., Bewley, G. C. & Sherald, A. F. *Genetics* **84**, 287-311 (1976).
16. Steller, H. & Pirrotta, V. *EMBO J.* **3**, 165-173 (1984).
17. Anderson, K. V. & Lengyel, J. *Dev. Biol.* **70**, 217-231 (1979).
18. Rubin, G. M. & Spradling, A. C. *Science* **218**, 348-353 (1982).
19. Lohs-Schardin, M., Cremer, C. & Nüsslein-Volhard, C. *Dev. Biol.* **73**, 239-255 (1979).
20. Lindsley, D. L. & Grell, E. H. *Publ. Carnegie Instn* **627** (1968).
21. Pirrotta, V., Hadfield, C. & Pretorius, G. H. *J. EMBO J.* **2**, 927-934 (1983).
22. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **24**, 5294-5299 (1979).
23. Jäckle, H. *Wilhelm Roux Arch. dev. Biol.* **188**, 225-233 (1980).
24. Thomas, P. S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201-5205 (1980).

LETTERS TO NATURE

Millisecond pulsars, gravitational waves and the inflationary universe

Lawrence M. Krauss

Lyman Laboratory of Physics, Harvard University,
Cambridge, Massachusetts 02138, USA

Signals from millisecond pulsars, being exceptionally accurate clocks, are sensitive to a cosmic background of gravitational waves¹⁻³. Present measurements imply that the energy density in such waves in an octave bandwidth near 10⁻⁷ Hz must be $\leq 10^{-3}$ of the closure density². Gravitational waves of these frequencies may be produced by a variety of sources of astrophysical or cosmological interest^{3,4}. I investigate here the sensitivity of a millisecond pulsar timing signal to the background of gravitational waves expected from an inflationary phase of expansion during the early Universe⁵. The spectrum of such waves, which is flat, extends up to and beyond periods of the order of the present Hubble time. As such, their amplitude can be constrained by measurements of the present quadrupole anisotropy of the black-body background⁶⁻⁸. I demonstrate here that the bound on inflation derived from millisecond pulsar timing measurements² is, at present, much weaker than this bound. However, the millisecond pulsar bound should, with new technology and the possible discovery of new pulsars, improve dramatically and it may eventually surpass the present quadrupole bound. It is in principle sensitive to the shape of the gravitational wave spectrum and can thus serve as a stronger test of inflationary predictions.⁹⁻¹¹ A positive millisecond pulsar signal can also be used to measure photon reheating processes in the early Universe below temperatures of the order of 100 MeV.

Starobinskii⁵ first demonstrated that a period of exponential expansion in the early Universe would result in a flat spectrum of gravitational waves today. At the time of horizon crossing, they have a fixed amplitude *A*, given by the relation,

$$A^2 = (2\pi)^{-2} [H/M_{\text{Pl}}]^2 \quad (1)$$

where *H* is the Hubble parameter during the phase of exponential expansion and *M_p* is the planck mass (1.2×10^{19} GeV).

Waves produced during inflation, which are only now entering the horizon, can dominate, through the Sachs-Wolfe effect, the large-scale quadrupole anisotropy of the microwave background. Several authors have used the limits on this anisotropy to bound the amplitude *A* and thus the Hubble constant of the inflationary

phase⁶⁻⁸. They find the bound $A \leq 0(10^{-4})$, implying that inflation takes place at a scale $\approx 4 \times 10^{17}$ GeV.

Using the relation $\langle T^{00} \rangle = k^2 A^2 / 8\pi G$ (where *k* is the physical wavenumber and where we have summed over the two helicity states) for gravitational waves, equation (1) leads to the result⁵,

$$\Omega_g(k)|_{\text{hc}} = 4/3 [H/M_{\text{Pl}}]^2 \quad (2)$$

where $\Omega_g(k)|_{\text{hc}}$ is the ratio to closure density of the energy density in gravitational waves of physical wavenumber *k* at horizon crossing.

The bound from the quadrupole anisotropy thus implies $\Omega_g(k) \leq 10^{-8}$ for gravitational waves now crossing the horizon. The energy density of waves now inside the horizon can be substantially less than this, owing to the effects of redshifting. Thus, the quadrupole bound is at best four orders of magnitude below that presently obtainable from the millisecond pulsar. However, two important factors make the millisecond pulsar constraints warrant further discussion. First, the pulsar sensitivity to a flat spectrum increases as the fourth power of the measuring time. Next, it is sensitive to waves with a vastly different history than those probed by the quadrupole bound and thus can yield important new information about several fundamental facets of standard cosmology.

First, I consider millisecond pulsar sensitivities with the idea of measuring a potentially-flat inflationary spectrum. The maximum delay in a signal caused by a passing gravitational wave of period *P* and amplitude *A* is of the order of *AP* (refs 1, 3). If the experiment has an inherent timing noise Δt then the smallest detectable *A* is $A \sim \Delta t/P$. Then, using the expression given earlier for the energy density of gravitational waves, pulsar timing measurements over a period $T \leq P$ are sensitive at best (within a factor of the order of one) to an energy density^{1,3,12}

$$\rho_g \geq (8\pi G)^{-1} (2\pi \Delta t)^2 P^{-4} \quad (3)$$

For gravitational waves with periods greater than the measuring time *T*, the sinusoidal time delay caused by the wave is harder to extract via polynomial fits to intrinsic time delays, owing to competing effects of spin down, and so on. By expanding this sinusoidal delay in power of *T/P*, only the third and higher terms can be distinguished from these unknown intrinsic pulsar features^{1,3}. This implies that for these longer wavelength waves the pulsar is sensitive to energy densities of the order of

$$\rho_g \geq (4G)^{-1} \Delta t^2 T^{-4} [P/T]^2 \quad (4)$$

For gravitational wave distributions with significant components of their energy density (that is, large amplitudes) in long wavelength modes, equation (4) may represent a greater sensitiv-

ity than equation (3). However, for a flat spectrum, equation (3) represents the optimal sensitivity.

Gravitational waves with periods today of $\leq 10^8$ yr entered the horizon during a period of radiation domination, according to the standard cosmological model. Having entered the horizon, the energy density of gravitational waves redshifts like radiation from its initial, wavenumber-independent value given in equation (2). If, as predicted by inflationary cosmology, the observed universe is flat to great precision, then Ω_{rad} was about unity at this time and, therefore, $\rho_g/\rho_{\text{rad}}|_{\text{hc}} \sim \Omega_g(k)|_{\text{hc}}$. Then, assuming no photon reheating, because gravitational waves have been redshifting like radiation, $\Omega_g(k)|_{\text{now}} \sim \Omega_g(k)|_{\text{hc}} \times \Omega_\gamma|_{\text{now}} \sim \Omega_g(k)|_{\text{hc}} \times 10^{-4}$ for waves which entered the horizon during the radiation dominated era. Using the present quadrupole limit, $\Omega_g|_{\text{hc}} \leq 10^{-8}$, we thus expect that $\Omega_g(k)|_{\text{now}} \leq 10^{-12}$ for these waves. Because under the assumption of no reheating their spectrum is flat, we can use equation (3) to compare the millisecond pulsar sensitivity with this limit.

The present pulsar limit on Ω_g is due to an observed timing stability of $\Delta_t \sim 10^{-6}$ s over a measuring period of several months². An order of magnitude improvement in measurement is expected soon because of improvements in Earth-based timekeeping technology (J. Taylor, personal communication). Thus, in principle, with a timing stability of 10^{-7} s and a measuring period of the order of 10 yr, the millisecond pulsar could yield a bound $\Omega_g(2\pi/10 \text{ light yr})|_{\text{now}} \leq 10^{-12}$, which is competitive with the present quadrupole limit.

The above discussion demonstrates what sensitivity is, in principle, achievable using the millisecond pulsar. Whether the limit quoted is practical is not at all clear. No clocks on Earth have such a good timing stability over such a long period. Also, other noise, both pulsar related and otherwise, can easily lead to greater timing instability in the signal. Of course, the discovery of other millisecond pulsars would greatly aid in handling such noise, and in improving timing measurements.

In deriving the millisecond pulsar bound above, we assumed there had been no reheating since horizon crossing. However, this is probably not the case for the waves of interest. The millisecond pulsar signal, unlike the quadrupole anisotropy, is sensitive to waves whose amplitude depends on the details of expansion since horizon crossing. (Of course, if the microwave background did not really represent a last scattering surface at $z = 1,000$, but instead was caused by a much more recent event, then the quadrupole anisotropy too can be dominated by waves which have long since entered the horizon.) Now, because of the effects of redshifting (with $P \sim t^{2/3}$ during the period of matter-dominated expansion since about recombination, and $\sim t^{1/2}$ before that time), waves from inflation with periods $\leq 10^3$ yr today entered the horizon before the time of nucleosynthesis, and those with a period of the order of 10 yr entered the horizon when the temperature was ~ 100 MeV ($t = 10^{-4}$ s). Even standard cosmology does not predict that the photon temperature has dropped according to the relation $T \sim R^{-1}$ since that time.

This raises the possibility that measurements made with the millisecond pulsar may give detailed information on the evolution of the Universe since very early times. In particular, there is no reason to expect that the spectrum is flat in the region of interest, nor that the amplitude of gravitational waves will be given simply by the product of their amplitude at horizon crossing and the simple redshift factor given above. If there has been any reheating (equilibrium, or out of equilibrium) since their entry in the horizon then their energy density is given today by,

$$\Omega_g(k)|_{\text{now}} = \Omega_g(k)|_{\text{hc}} \times \Omega_\gamma|_{\text{now}} \times [\Pi\{T_i/T_f\}^4 \times \Pi\{S_{\text{rad}}/S_{\text{rad}}'\}^{4/3}] \quad (5)$$

where T_i , T_f are the temperatures before and after each non-adiabatic reheating and S_{rad} and S_{rad}' are the entropy densities of radiation before and after each adiabatic heating due to particle-antiparticle annihilation.

Eventual detection of gravitational waves by way of the millisecond pulsar could resolve these reheating process under two

different circumstances. First, if the pulsar results were combined with a positive quadrupole anisotropy determination then the factor in square brackets in equation (4) could immediately be estimated. Even if there were no quadrupole measurement to use, however, departures from a flat spectrum might be compared with expected deviations due to reheating to check whether the waves were primordial in origin.

What kind of reheating might we then expect to detect? Minimally, in the standard cosmological model, it is assumed that the expansion has essentially always been adiabatic. In this case, adiabatic heating occurs as the temperature falls below the mass thresholds for various known particles, whose number density is subsequently reduced by annihilation and whose entropy is redistributed to the remaining particles. Thus, for temperatures below about 100 MeV, we expect two such periods, as first muons and then electrons annihilate. The first process reduces the ratio of the energy density of gravitational waves to that of photons by a factor $(11/18)^{4/3}$, and the second by the factor $(4/11)^{4/3}$, thus resulting in a net decrease in Ω_g for waves of period < 10 yr by about a factor of 7, compared with our earlier estimate.

Of course, the adiabatic expansion picture may not be correct back to temperatures of 100 MeV. Although there are some constraints on changes in entropy since nucleosynthesis¹³, the period from ~ 1 to 100 MeV is not empirically constrained. In particular, it is a property of several emerging particle physics models, notably supersymmetry, that there would tend to be out-of-equilibrium decay of new heavy particles in the early Universe. Indeed, some of the constraints on these models come from requiring that these decays must either occur before nucleosynthesis, or reheat the temperature to above 1 MeV (refs 14, 15). It would be useful for model building if some actual empirical information was available about entropy generation during this period. Finally, although nucleosynthesis calculations imply that there was no significant reheating since that time, it has been pointed out¹⁶ that there is room for flexibility, especially when one takes into account the fact that particle decays can change the remnant deuterium and helium abundances through destruction of these elements¹⁷. A direct empirical measurement would thus be very useful, both in checking standard nucleosynthesis models, and in solidifying the various constraints on particle physics (see, for example, refs 13, 18) which have been based on them.

In conclusion then, the millisecond pulsar is, at present, less powerful than the quadrupole anisotropy at constraining inflation. However, with improved technology, it should exceed present quadrupole bounds and in the process not only give the best limits on the scale of inflation, but also provide, in principle, empirical information on reheating in the early Universe, the importance of which, to particle physics and cosmology, cannot be overemphasized.

I thank J. Taylor for his Loeb Lectures which first interested me in this subject, E. Witten, S. Coleman, C. Hogan and M. Wise for helpful discussions and the Aspen Center for Physics, where this work was completed. L.M.K. is a junior fellow, Harvard Society of Fellows. Research was supported in part by the NSF under contract no. PHY-82-15249.

Received 23 July; accepted 30 October 1984.

- Bertotti, B., Carr, B. J. & Rees, M. J. *Mon. Not. R. astr. Soc.* **203**, 945-954 (1983).
- Taylor, J. Preprint, *Pulsar Timing, Millisecond Pulsars, and Gravitational Waves* (Princeton University, 1984).
- Witten, E. *Phys. Rev. D* **30**, 272-285 (1984).
- Turok, N. Santa Barbara Preprint TH-3 (1984).
- Starobinskii, A. A. *JETP Lett.* **30**, 682-685 (1979).
- Rubakov, V. A., Sazhin, M. V. & Veraskin, A. V. *Phys. Lett.* **115B**, 189-192 (1982).
- Fabbri, R. & Pollock, M. D. *Phys. Lett.* **125B**, 445-448 (1983).
- Abbott, L. F. & Wise, M. B. Cal-Tech Preprint, CALT-68-1100 (1984).
- Guth, A. *Phys. Rev. D* **23**, 347-360 (1981).
- Linde, A. D. *Phys. Lett.* **108B**, 389-392 (1982).
- Albrecht, A. & Steinhardt, P. J. *Phys. Rev. Lett.* **48**, 1220-1224 (1982).
- Detweiler, S. *Astrophys. J.* **234**, 1100-1107 (1979).
- Yang, J., Turner, M. S., Steigman, G., Schramm, D. N. & Olive, K. *Astrophys. J.* (in the press).
- Weinberg, S. *Phys. Rev. Lett.* **48**, 569-573 (1982).
- Krauss, L. M. *Nucl. Phys. B* **229**, 556-567 (1983).
- McKellar, B. H. J. & Pavasa, S. University of Hawaii Preprint UH-511-536-84.
- Lindley, D. *Mon. Not. R. astr. Soc.* **188**, 15P-19P (1979).
- Krauss, L. M. Harvard University Preprint, HUTP-84/A050 (1984).

Multiple hotspots in extragalactic radio sources

A. G. Williams & S. F. Gull

Mullard Radio Astronomy Observatory, Cavendish Laboratory,
Madingley Road, Cambridge CB3 0HE, UK

Recent observations with sub-kpc resolution show that the hotspots in powerful extragalactic radio sources (for example, Cygnus A¹, 3C351², 3C268.4³, 3C133⁴, 3C20⁴, 3C196⁴) often exhibit internal double structure. One subcomponent is usually extremely compact, of size <1 kpc, whereas the other, which is more diffuse, often appears to point away from the smaller component and is usually further from the nucleus. We present here a model of a jet with variable direction that is able to account for this structure, interpreting the compact subcomponent as the initial point of impact of the jet with the intergalactic medium and the more diffuse subcomponent as a 'splatter-spot' formed by jet material deflected asymmetrically by the compact subcomponent.

In the context of the now-accepted jet model of radio sources^{5,6}, authors differ in their interpretation of where the jet first deposits its energy. Laing⁴ has suggested that the compact subcomponent is the current point of impact of the jet on the surrounding medium whereas the diffuse subcomponent either marks a previous hotspot created at a time when the direction of the jet was different, or is produced by material escaping from the compact hotspot (this model is also discussed by Lonsdale and Barthel⁷). Conversely, Kronberg and Jones⁸ propose that the diffuse subcomponent occurs at the end of the jet and that the compact subcomponent might arise as a short-lived instability in the contact surface between the lobe and the shocked intergalactic medium (IGM). Begelman *et al.*⁹ suggest that the subcomponents may either result from jitter in the orientation of the jet or be simultaneous points of impact of a jet that has become subdivided, whereas Norman *et al.*¹⁰ propose that the double hotspot structure may be interpreted as the signature of the triple-shock configuration found at the working surface of axisymmetric jets. Previous numerical simulations of hotspots^{10,11} have concentrated on axisymmetric models predicting a single working surface characterized by a Mach disk¹⁰. In this paper, we describe the result of fully three-dimensional calculations in which the direction of the jet is varied.

We performed the numerical simulations by integrating the non-relativistic equations of ideal compressible flow, as described by us¹². The hydrodynamic code is a 'flux-corrected'¹³ 'fluid-in-cell'¹⁴ scheme and was carried out on an (X, Y, Z) grid of $120 \times 120 \times 12 = 172,800$ cells on the Starlink VAX 11/780 computer at Cambridge. We stretched the cells in two of the directions to provide a domain $120 \times 360 \times 48$ (so that the computing cells are rectangular parallelepipeds with sides $1 \times 3 \times 4$). The jet was introduced in the Y-direction with a circular cross-section and radius 8 units, corresponding to eight zones in the X-direction and two zones in the Z-direction. Elsewhere, the fluid was allowed to escape freely from the computing grid, except in the $Y=0$ and $Z=0$ planes where reflective boundary conditions were used. Because of the computational cost of these simulations, we could only investigate a limited region of the available parameter space, which we chose on the basis of present knowledge of radio sources. Axisymmetric simulations suggest that jets in sources of Fanaroff and Riley (FR) class II¹⁵ must be light and supersonic because only this type exhibits hotspots and broad lobes. For these simulations we chose a density ratio of 1:10 because simulations of extremely diffuse jets with density ratios of 1:1,000 are computationally inefficient and at present impracticable in three dimensions. Our experience in two dimensions suggests that a well-developed jet with a density ratio of 1:10 has the same essential features (for example, hotspots, cocoon, backflow) as still lighter jets. To isolate the effect of breaking axisymmetry, we fixed the input

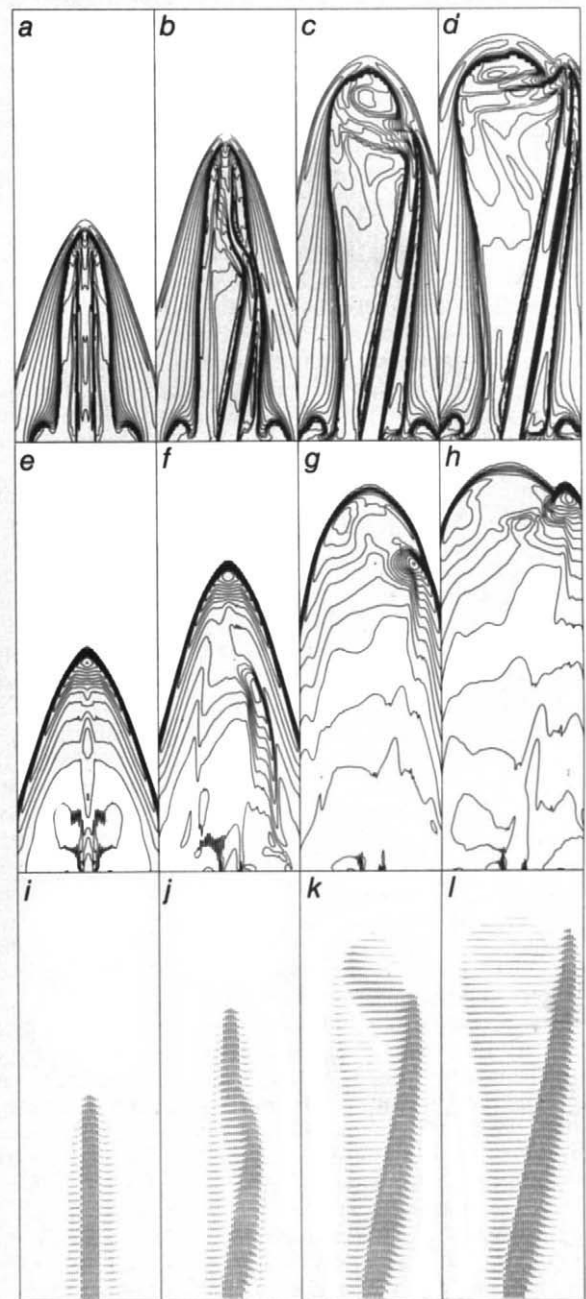


Fig. 1 The distributions of density (a-d), pressure (e-h) and projected velocity vectors (i-l) (at every other grid point) in the $Z=0$ plane for the 10° simulation shown at: $t=14.1$, $t=19.4$, $t=27.7$ and $t=33.1$. The density contours are logarithmic with contouring ratio 1.4:1. The pressure contours are logarithmic with contouring ratio 1.3:1.

Mach number of the jet equal to 10 (relative to its internal sound speed) and imposed thermal pressure balance between the unshocked jet and IGM. In any case the variation in morphology with Mach number and pressure ratio is known to be weak provided that the jets are sufficiently supersonic.

We investigated departure from axisymmetry by varying the jet direction suddenly. Although slow precession of the primary collimator¹⁶ is a plausible cause of asymmetry in astrophysical jets, we did not attempt to model this in detail because of the extra computational cost of providing equal resolution in all three directions. Many of the features that would result from precession can nevertheless be seen in our models; different rates of precession correspond to the different epochs in our simulations and variations in the cone angle are similar to different abrupt changes in the jet axis.

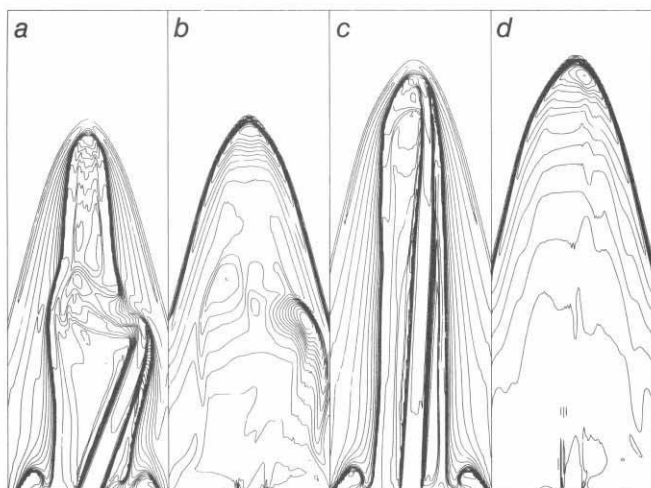
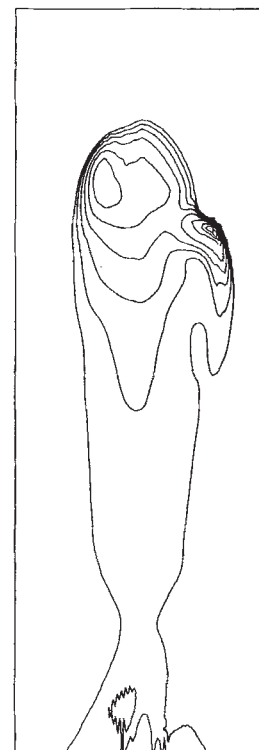


Fig. 2 The distribution of density and pressure in the $Z=0$ plane for: *a* and *b*, the 20° simulation at $t=21.5$; *c* and *d*, the 3° simulation at $t=25.2$. The contour levels are the same as in Fig. 1.

It is natural to compare the results of these simulations with axisymmetric calculations. Figure 1*a* and *e* shows the distribution of density and pressure in the $Z=0$ plane just before the direction of the jet is altered. These simulations and others carried out with higher resolution confirm that the general features previously reported^{10,11} are not severely affected by artefacts due to the cylindrical computing mesh. Nevertheless, a subtle but important difference is that the string of on-axis compressions and rarefactions, caused by the interaction of the jet with its backflow, are much stronger in the axisymmetric case (compare Fig. 1*e* and *g*). We interpret this as an artificial 'refocusing', in the strictly axisymmetric case, which is unlikely to be found in real sources. Attempts to model the knots in jets using axisymmetric simulations¹⁰ may consequently have overestimated the strength of knots that could be produced. However, because of the decrease in resolution in the Y and Z directions, the internal shocks and the jet-shock in our simulations are broader than the nearly plane bow-shock, which has no room to diffuse in the directions perpendicular to the shock.

In our first simulation, the direction of the jet is changed suddenly by 10° . The jet immediately encounters material that is not receding from it and forms a weak warm-spot (we associate the regions of high pressure occurring behind shocks with enhanced radio emission¹⁷) which quickly crosses the low-density cocoon. When the jet reaches the contact surface, the warm-spot pressure rises sharply as the jet then encounters material denser than itself to which it gives up a large fraction of its kinetic energy. An axisymmetric jet reverses its direction at the hotspot by terminating in a Mach disk¹⁰, with a normal shock across the centre of the jet and conical 'reflected shocks' at the edge of the jet deflecting material away from the axis. In contrast, a non-axisymmetric jet adjusts to the asymmetric presence of the cocoon wall by producing an oblique compression wave (Fig. 1*b*, *f*). After a small distance, this wave coalesces to form an oblique shock, although, in our simulations, we cannot resolve the exact point at which this occurs. The strength of the shock is zero where it first forms, and increases with distance from the cocoon wall¹⁸. After passing through the shock, the flow is bent by a rarefaction wave which increases the Mach number of the jet. We note that, for these flow parameters, the flow remains supersonic during its passage through the hotspot. The jet then recrosses the cocoon (which still contains less-dense material), meeting the opposite wall at a secondary hotspot or 'splatter-spot' (Fig. 1*c*, *g*). In its early stages, the secondary jet-shock is sufficiently oblique for the jet to be reflected supersonically, forming yet another hotspot near the apex of the cocoon; at a later stage, the secondary hotspot becomes less oblique and begins to resemble the Mach disk of axisymmetric simulations with subsonic outflow (Fig. 1*d*, *h*), whereas later

Fig. 3 An indication of the location of radio emission for the 10° simulation at $t=27.7$, calculated simply by integrating the pressure inside the contact surface along a line of sight through the source. The contour levels are equally spaced.



still, the primary hotspot dominates the dynamics, producing its own axisymmetric backflow which cuts off the flow to the secondary hotspot. The distribution of density, in cross-sections through the jet, shows that the asymmetric hotspot cuts a narrow channel through the IGM because, unlike axisymmetric jets, its waste products are able to escape asymmetrically into the previously inflated cocoon.

We next consider the question of hotspot compactness. The primary hotspot (where the asymmetric jet first meets the cocoon wall) is initially elongated parallel to the jet-shock, although, as the distortion in the contact surface increases, the jet-shock becomes less oblique and the hotspot correspondingly more compact. In the plane perpendicular to the jet-shock, the primary hotspot is of similar size to the width of the jet before it enters the hotspot region. By comparison, the secondary hotspot (formed at the second impact with the walls of the cocoon) is much larger. Three factors seem relevant: (1) the jet is genuinely more diffuse as a result of the expansion wave; (2) the direction of the reflected jet varies as the primary hotspot distorts the cocoon wall; (3) the reflected jet can expand in the Z -direction inside the previously inflated cocoon. Furthermore, the primary hotspot of a precessing jet may be the site of a Rayleigh-Taylor instability, as the contact surface will accelerate in the direction of the denser IGM. Figure 1*c* shows a smearing of the contact surface suggesting that this may indeed be occurring.

Density and pressure contours in the $Z=0$ plane for the simulation in which the jet angle was changed by 20° are shown in Fig. 2*a* and *b* and are broadly similar to those in the 10° simulation. The primary and secondary hotspots are both visible inside the outer bow-shock. At the primary hotspot, the post-jet-shock material is again able to flow away from the central engine. Figure 2*b* shows that the hotspot is not edge-brightened on this side—the reason being that a smaller fraction of its kinetic energy is given up to the less-dense material in the cavity; also noticeable is a pressure valley between the primary and secondary hotspots. The simulation (Fig. 2*c* and *d*) in which the jet angle was changed by 3° is, however, qualitatively different from the other two: the jet is deflected by the wall of the cocoon but it arrives at the apex of the cavity with its kinetic energy essentially intact and forms a single asymmetric hotspot. Thus no 'splatter-spot' results; instead the post-hotspot material is deflected through 180° producing asymmetric backflow.

We have shown above that a simple single-jet fluid-dynamical model is able to explain the phenomenon of multiple hotspots without any need for additional mechanisms such as segmentation of the jet, special viewing angles or localized instabilities. Below a critical angle of jet deflection, a single distorted hotspot with asymmetric backflow forms at the apex of the cocoon. This structure is reminiscent of that seen in NGC315 (ref. 19). Above the critical angle, a compact subcomponent forms where the jet first meets the cocoon wall: more diffuse subcomponents then result, as 'splatter-spots', from the oblique shock formed there. Our fluid-dynamical model makes the following predictions about these sources.

(1) One subcomponent will be much more compact than the others. This is already known to be the case in many sources.

(2) The jet from the nucleus currently points at the compact subcomponent. Although present observations have insufficient dynamic range to determine the path of the jet in most sources, the jet in 3C133 (ref. 4) certainly appears to point towards the compact subcomponent.

(3) We expect that the compact subcomponent will be curved and elongated because the jet interacts obliquely with the cocoon wall. The hotspot should decrease in strength towards the outside of the bend corresponding to the decrease in strength of the jet-shock there. The precise direction of elongation depends on the relative strengths of the jet-shock and the rarefaction wave. The hotspots in Cygnus A (ref. 1) and 3C351N (ref. 2) both show extensions of this type.

(4) The usual form of jet-shock in astrophysical jets will be oblique, with the well-studied Mach disk only arising as a special case appropriate near axisymmetry.

(5) Subsequent subcomponents have jet-shocks that are nearer to normal incidence because the Mach number is reduced at each hotspot. Powerful (high Mach number) jets produce multiple hotspots whereas weaker sources have subsonic outflow from the primary hotspot. If any of the shocks are almost normal, then no further hotspots are possible, as a shock inclined at 45° or more to the flow reduces the Mach number to order unity irrespective of its value upstream.

(6) Adiabatic loss in the rarefaction wave following each subcomponent may explain why few ridges of emission are found between multiple hotspots.

(7) The axial ratio of the cocoon is smaller than would be predicted by axisymmetric models, because its sideways expansion is enhanced by the inherent variation in the jet direction and by the formation of 'splatter-spots'.

(8) In our interpretation of a precessing jet, the compact subcomponent moves around the head of the cocoon as envisaged in Scheuer's 'Dentist's Drill' model²⁰. The time-averaged hotspot advance speed should consequently be scaled down by the ratio of the size of the subcomponent to its perpendicular distance from the axis of rotation. Thus the advance speeds of the compact and diffuse subcomponents may be comparable, even though the compact subcomponent has a higher energy density.

(9) Radio emission is confined solely to the region inside the contact surface (Fig. 3). If the shocked IGM, behind the bow-shock, contained a significant number of radio-emitting particles, then the structure of such a source would be qualitatively different from that observed. In particular, the hotspots would be surrounded by edge-brightened emission from particles accelerated at the very broad bow-shock. Also, the distortions observed in radio-trail sources¹² and in the bridges of classical doubles²¹ are not consistent with emission from the predominantly convex bow-shock.

We thank Drs Robert Laing and Peter Scheuer for helping us to formulate these ideas.

5. Blandford, R. D. & Rees, M. J. *Mon. Not. R. astr. Soc.* **169**, 395-415 (1974).
6. Scheuer, P. A. G. *Mon. Not. R. astr. Soc.* **166**, 513-528 (1974).
7. Lonsdale, C. J. & Barthel, P. D. *Astr. Astrophys.* **135**, 45-52 (1984).
8. Kronberg, P. P. & Jones, T. W. *IAU Symp.* No. 97 (eds Heeschen, D. S. & Wade, C. M.) 157-160 (Reidel, Dordrecht, 1982).
9. Begelman, M. C., Blandford, R. D. & Rees, M. J. *Rev. Mod. Phys.* **56**, 255-351 (1984).
10. Norman, M. L., Smarr, L., Winkler, K.-H. A. & Smith, M. D. *Astr. Astrophys.* **113**, 285-302 (1982).
11. Wilson, M. J. & Scheuer, P. A. G. *Mon. Not. R. astr. Soc.* **205**, 449-463 (1983).
12. Williams, A. G. & Gull, S. F. *Nature* **310**, 33-36 (1984).
13. Book, D. L., Boris, J. P. & Hain, K. J. *computational Phys.* **18**, 248-283 (1975).
14. Gentry, R. A., Martin, R. E. & Daly, B. J. *J. computational Phys.* **1**, 87-118 (1966).
15. Fanaroff, B. L. & Riley, J. M. *Mon. Not. R. astr. Soc.* **167**, 31P-35P (1974).
16. Ekers, R. D., Fant, R., Lari, C. & Parma, P. *Nature* **276**, 588-590 (1978).
17. Bell, A. R. *Mon. Not. R. astr. Soc.* **182**, 147-156 (1978).
18. Courant, R. & Friedrichs, K. O. *Supersonic Flow and Shock Waves*, 294-297 (Interscience, New York, 1948).
19. Willis, A. G., Strom, R. G., Bridle, A. H. & Fomalont, E. B. *Astr. Astrophys.* **95**, 250-265 (1981).
20. Scheuer, P. A. G. *IAU Symp.* No. 97 (eds Heeschen, D. S. & Wade, C. M.) 163-165 (Reidel, Dordrecht, 1982).
21. Leahy, J. P. & Williams, A. G. *Mon. Not. R. astr. Soc.* **210**, 929-951 (1984).

Periodic comet showers and planet X

Daniel P. Whitmire & John J. Matese

Department of Physics, PO Box 44210,
University of Southwestern Louisiana, Lafayette,
Louisiana 70504, USA

The discovery¹ that Pluto's mass is insufficient to explain the discrepancies in the motions of the outer planets has led to the prediction of a tenth planet (planet X) of mass $\sim 1-5 M_{\oplus}$ beyond the orbit of Pluto (50-100 AU)^{2,3}. Further, the existence of a belt or disk of comets beyond the orbit of Neptune (≥ 35 AU) has been proposed in connection with some theories of the origin of the Solar System^{4,5} as a possible source of short-period comets⁶ and indirectly as a source of long-period comets^{7,8}. Here we point out that the existence of both planet X and the comet disk at their expected distances may explain not only the observed planetary motions and the origin of comets, but also the recently reported 28-Myr periodicity in terrestrial cratering⁹⁻¹¹ and in mass extinctions^{11,12}. The cratering period is associated with the precession of the perihelion of planet X caused by the perturbations of the outer planets.

During the lifetime of the Solar System, planet X will have swept out a gap in the comet disk of width determined by its perihelion and aphelion distances⁴. Assuming the planet's orbit has modest eccentricity and inclination to the planetary plane, then the perihelion (aphelion) will pass close to the inner (outer) gap edge twice during each 56-Myr precession period. The resulting shower of comets from either or both edges will be concentrated near the planetary plane, thereby facilitating capture into short-period (SP)-comet orbits. In the steady state, comets near the gap edges will be resupplied between showers by diffusion⁶.

The precession period, T_{ω} , of planet X can be approximated by considering the planets as uniform rings of masses M_p and radii equal to their semi-major axes a_p . The perturbing potential (and precession rate) on a planet of semi-major axis a outside a ring is well known (see, for example, ref. 13). Using these results we obtain for the semi-major axis

$$a = \left[\frac{3}{2} T_{\omega} \left(\sum_p M_p a_p^2 \right) \frac{|1 - \frac{5}{4} \sin^2 i|}{(1 - e^2)^2} \right]^{2/7} \quad (1)$$

where the masses are in solar units, semi-major axes are in AU and T_{ω} is in yr. Here i and e are respectively the orbital inclination and eccentricity of planet X. The required semi-major axis is seen to be quite insensitive to these orbital elements and the precise value of T_{ω} . Setting $T_{\omega} = 56$ Myr gives $a \approx 100$ AU. At this distance planet X is potentially observable both optically^{2,3} and in the infrared¹⁴. The infrared astronomy satellite (IRAS) 100- μ m band is capable of detecting a $1 M_{\oplus}$ planet at ~ 160 AU¹⁴.

Received 8 September; accepted 24 October 1984.

1. Alexander, P., Brown, M. T. & Scott, P. F. *Mon. Not. R. astr. Soc.* **209**, 851-868 (1984).
2. Kronberg, P. P., Clarke, J. N. & van den Bergh, S. *Astr. J.* **85**, 973-980 (1980).
3. Lonsdale, C. J. & Morison, I. *Mon. Not. R. astr. Soc.* **203**, 833-851 (1983).
4. Laing, R. A. *IAU Symp.* No. 97 (eds Heeschen, D. S. & Wade, C. M.) 161-162 (Reidel, Dordrecht, 1982).

In our estimate of a , we have neglected any contribution due to the comet disk itself. The mass of the inner disk can be limited to $\sim 1 M_{\oplus}$ on the basis of the absence of any observable perturbations of the orbits of several periodic comets¹⁵ and the IRAS detection limits¹⁶. If the inner-disk mass is $\leq M_{\oplus}$, a correction for this would increase a by less than a few per cent. A massive outer disk¹⁷ could conceivably be primarily responsible for the precession period; further, if planet X (like Pluto) is in a resonant orbit with Neptune, then equation (1) may not necessarily be a good approximation—in either case the perihelion will still precess or librate, but the semi-major axis required for a 56-Myr periodicity will in general differ from the above estimate.

In the context of our model the inclination of the orbit of planet X should be greater than the angular thickness of the comet disk near the edge. The mean inclination of the SP comets (assuming they originate in the disk) suggests that $i \geq 10^\circ$. Thus both the semi-major axis and inclination required by the precession model are consistent with previous predictions based on the anomalous residuals in the motions of Neptune and other planets^{2,3}.

As it is unlikely that comets were formed in the low-density environment of the Oort cloud, it is often assumed that they originated in the neighbourhood of Uranus and Neptune⁴⁻⁸. Planetary perturbations would then populate a hypothetical^{18,19} massive inner cloud ($\approx 2 \times 10^4$ AU) while subsequent stellar perturbations could maintain the observed Oort cloud^{7,8}. In this picture we would expect a primordial belt or disk of comets beyond Neptune's influence (≥ 35 AU) still to exist⁴⁻⁸. The presence of Pluto, whose mass is now known to be $\sim 10^{-3} M_{\oplus}$, would, unlike planet X, not significantly perturb the disk.

Because of apparent difficulties in capturing a sufficient number of SP comets from long-period comets, Fernandez⁶ has proposed a model in which the observed SP comets originate in such a disk. Comets from the inner-disk edge are passed on to Neptune by a diffusion process driven by massive comets within the disk. According to Fernandez, about one comet per year with mass $\geq 10^{15}$ g is needed to diffuse into Neptune's zone of influence to account for the steady-state flux. The comets are then passed on to Uranus and Saturn, finally becoming Jupiter-family SP comets. The process is highly efficient with about half the comets being passed on by each planet so that $\sim 1/16$ of all comets diffusing into Neptune's zone end up as SP comets. In the context of a diffusion model, conservation of angular momentum requires that a roughly equal flux of comets diffuse outwards, which is analogous to an isolated accretion disk whose outer rings must expand so that the inner rings can accrete. In the following analysis we assume only that the steady-state SP comets originate in the comet disk and that they diffuse into Neptune's influence at the 'observed' rate of ~ 1 comet yr^{-1} . The diffusion may be driven by massive comets or by some other mechanism.

We shall assume that the diffusion process continuously repopulates transition zones of scale length, σ , near the gap edges. If the gap is determined by the perihelion, q , and aphelion of planet X, then a steady state will exist in which the number of comets perturbed out of the transition zones every 28 Myr by planet X equals the number of comets diffusing into the zones during this interval. To model the time dependency of the flux of scattered comets, we assume a gaussian distribution of comets in the transition zones. The flux of comets scattered from the inner-gap edge is then approximated by

$$\phi_{\text{scattered}}(t) \approx A \phi_{\text{diffusion}} \exp \left[- \left(\frac{r_c - q}{\sigma} \right)^2 \right] \quad (2)$$

where r_c is the distance of closest approach of planet X to the disk during a single orbital period and A is a normalizing factor. When the argument of perihelion, ω , is near 0 or π it can be shown that for a thin disk

$$r_c - q \approx \frac{2\pi^2 e(1-e)}{(1+e)} a \left(\frac{t}{T_\omega} \right)^2 \quad (3)$$

Inserting this into equation (2) and evaluating A from the

average steady-state condition $\int_0^{1/2 T_\omega} \phi_{\text{scattered}} dt = \frac{1}{2} T_\omega \phi_{\text{diffusion}}$, gives

$$\frac{\phi_{\text{scattered}}(t)}{\phi_{\text{diffusion}}} \approx 1.2 K^{1/2} \exp \left[-K^2 \left(\frac{t}{T_{\text{ext}}} \right)^4 \right] \quad (4)$$

where the extinction period $T_{\text{ext}} = \frac{1}{2} T_\omega = 28$ Myr and

$$K = \frac{\pi^2}{2} \frac{a}{\sigma} \frac{e(1-e)}{(1+e)} \quad (5)$$

We define the shower interval, Δt , as twice the time required for the flux to decrease from maximum ($= \phi_{\text{scattered}}(0)$) to $1/e$ of maximum, thus

$$\frac{\Delta t}{T_{\text{ext}}} \approx \frac{1}{11} \left[\frac{\sigma(\text{AU})(1+e)}{e(1-e)} \right]^{1/2} \geq \frac{2}{9} \sigma^{1/2} \quad (6)$$

In the diffusion model⁶ σ is typically ~ 0.3 AU depending on the maximum mass of the comets in the disk. However, the disk may be truncated by planet X itself, analogous to the truncation of stellar accretion disks by low-mass binary companions²⁰. This mechanism can be especially effective if resonances are important (see below). Thus, although $\sigma = 0.3$ AU gives an acceptable shower duration (for $0.2 \leq e \leq 0.5$), values much less than this may also be possible.

We now discuss the number of terrestrial impacts of comets of mass $\geq 10^{15}$ g expected, given that 28×10^6 such comets are perturbed in the planetary plane during a shower. The fraction of these comets that are scattered into orbits with perihelia within the influence of Neptune is $\sim (2a_N/q)(1 - a_N/2q)$ where a_N is Neptune's semi-major axis¹⁸: setting $q = a(1-e) \approx 70$ AU gives $\sim 1/2$; these comets are passed on to Uranus, Saturn and Jupiter with a similar efficiency of $\sim 1/2$ ^{6,21}. The corresponding fraction of comets reaching Neptune's orbit that ultimately become part of the Jupiter family of SP comets is $\sim 1/16$ and thus the number of SP-shower comets is $\sim 28 \times 10^6 \times 1/32 \sim 10^6$. Of these Jupiter-family comets, ~ 0.1 will have perihelia ≤ 1 AU²², yielding $\sim 10^5$ Earth-crossers. The probability per perihelion passage that one of these SP comets will strike the Earth is $\sim 7 \times 10^{-9}$. Thus the expected number of terrestrial impacts per shower is $\sim (10^5 \text{ comets}) \times (n \text{ revolutions per comet}) \times (7 \times 10^{-9} \text{ impacts per revolution}) \approx 7 \times 10^{-4} n$. To restrict n , we require that the cratering rate due to SP shower and steady state comets be less than the total observed rate.

According to Weissman²³, average SP comets of mass $\geq 10^{15}$ g will produce ≥ 20 -km diameter craters. The observed mean cratering rate over the last 300 Myr, for craters of diameter ≥ 20 km, is $0.35 (\pm 0.13) \times 10^{-14} \text{ km}^{-2} \text{ yr}^{-1}$ corresponding to 50 impacts per 28 Myr interval. Assuming a mean mass index⁶ of 1.75, one of these impactors will have a mass $\geq 2 \times 10^{17}$ g. In the context of our basic model we expect the SP shower and steady-state contributions to cratering to be comparable. We thus require $2 \times 7 \times 10^{-4} n < 50$ or $n < 3.6 \times 10^4$ revolutions, which is compatible with various estimates of the physical lifetime of SP comets^{22,23,25}. The lower limit on n depends on the unknown number and mass of shower impacts necessary to trigger a mass extinction event.

Although our basic model implies comparable numbers of shower and steady-state impacts, the shower contribution might actually dominate as we have neglected the contribution of showers from the outer gap (aphelion). Further, we have neglected any contribution associated with resonant scattering, the potential importance of which is apparent in the Kirkwood gaps in the Asteroid belt and in Saturn's ring system. The contribution from the 2:1 and 3:2 resonances, in particular, may dominate the non-resonance scattering considered above²⁰. For $a \approx 100$ AU, these resonances would scatter comets in the inner disk near rings at 63 and 76 AU. Resonant scattering will be extremely effective if one of these rings is at or near perihelion, corresponding to $e \approx 0.37$ or 0.24 respectively. Resonant scattering at aphelion could also be important in the outer disk. These considerations suggest that the total number of shower comets may significantly exceed the number of steady-state SP comets. The

fraction of terrestrial craters due to the steady-state population of SP comets (including extinct cores) is highly uncertain (5–100%) owing to the uncertainty in the physical lifetime of these objects²³.

Random impacts by 10-km bodies are expected to occur with a mean frequency of ~ 50 Myr²⁶. The absence¹¹ of any associated random mass extinctions and the evidence that some extinctions take place over ~ 1 -Myr intervals both suggest that these events may be primarily associated with the enhanced rate of impacts during a shower (or some other shower-related phenomenon) rather than with the impact of a single 10-km comet, although such an event is also expected occasionally.

In conclusion, we emphasize that both planet X and the comet

disk have been previously postulated for reasons that are unrelated to periodic comet showers and mass extinctions. If both do indeed exist, then periodic comet showers at the observed frequency may be a natural consequence. The perihelion precession period of planet X is expected to be highly stable over a 250-Myr interval, in contrast to the solar companion model^{27,28} in which the orbital period is expected to random-walk away by $\sim 15\%$ over this same interval^{29,30}. At present, the fossil record cannot alone discriminate between these two predictions.

We thank Drs A. A. Jackson, M. E. Bailey and J. A. Fernandez for helpful suggestions. This work was supported in part by the NSF grant AST-8405077.

Received 27 July; accepted 24 October 1984.

- Harrington, R. S. & Christy, J. W. *Astr. J.* **86**, 442–443 (1981).
- Van Flandern, T. C., Kaplan, G. H., Plukkinen, K. F., Santoro, E. J., & Seidelmann, P. K. *Bull. Am. Astr. Soc.* **12**, 830 (1980).
- Abell, G. O. *Exploration of the Universe*, 349 (Saunders, Philadelphia, 1982).
- Kuiper, G. P. *Astrophysics* (ed. Hynek, J. A.) (McGraw-Hill, New York, 1951).
- Whipple, F. L. *The Motion, Evolution and Origins of Comets*, IAU Symp **45**, 401–408 (1972).
- Fernandez, J. A. *Mon. Not. R. astr. Soc.* **192**, 481–491 (1980).
- Shoemaker, E. M. & Wolfe, R. F. Preprint (paper presented at the Fifteenth Lunar and Planetary Science Conference, 1984).
- Fernandez, J. A. Preprint (paper presented at IAU Colloquium 83, 1984).
- Alvarez, W. & Muller, R. A. *Nature* **308**, 718–720 (1984).
- Rampino, M. R. & Strothers, R. B. *Nature* **308**, 709–712 (1984).
- Sepkoski, J. J. & Raup, D. M. in *Dynamics of Extinction* (ed. Elliott, D.) (Wiley, New York, in the press).
- Raup, D. M. & Sepkoski, J. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 801–805 (1984).
- Roy, A. E. *Orbital Motion*, 285 (Wiley, New York, 1978).

- Reynolds, R. T., Tarter, J. C. & Walker, R. G. *Icarus* **44**, 772–779 (1980).
- Hamid, S. E., Marsden, B. G. & Whipple, F. L. *Astr. J.* **73**, 727–729 (1968).
- Bailey, M. E. *Mon. Not. R. astr. Soc.* **205**, 47P–52P (1983).
- Bailey, M. E. *Nature* **302**, 399–400 (1983).
- Hills, J. G. *Astr. J.* **86**, 1730–1740 (1981).
- Bailey, M. E. *Mon. Not. R. astr. Soc.* **204**, 603–633 (1983).
- Lin, D. N. C. & Papaloizou, J. *Mon. Not. R. astr. Soc.* **186**, 799–812 (1979).
- Everhart, E. *Comets-Meteorites-Asteroids: Interrelations, Evolution and Origin* (ed. Del-somme, A. H.) 99–104 (University of Toledo, Ohio, 1977).
- Fernandez, J. A. *Astr. Astrophys.* **135**, 129–134 (1984).
- Weissman, P. R. *Geol. Soc. Am. spec. Pap.* **190**, 15–24 (1982).
- Grieve, R. A. F. & Dence, M. R. *Icarus* **38**, 230–242 (1979).
- Shoemaker, E. M. & Wolfe, R. F. in *The Satellites of Jupiter* (eds Morrison, D. & Matthews, M. S.) 277–339 (University of Arizona Press, Tucson, 1982).
- Wetherill, G. W. & Shoemaker, E. M. *Geol. Soc. Am. spec. Pap.* **190** 1–13 (1982).
- Whitmire, D. P. & Jackson, A. A. *Nature* **308**, 713–715 (1984).
- Davis, M., Hut, P. & Muller, R. *Nature* **308**, 715–717 (1984).
- Hut, P. *Nature* **311**, 638–641 (1984).
- Hills, J. *Nature* **311**, 636–638 (1984).

Ring structure in Saturn's outer magnetosphere

H. S. Mahra, A. K. Pandey, Vijay Mohan & B. B. Sanwal

Uttar Pradesh State Observatory, Manora Peak, Naini Tal, India

The presence of a particulate or gaseous ring in Saturn's outer magnetosphere at distances of 14 and 19 radii from Saturn was predicted by Lazarus *et al.*¹ based on the analysis of ion density measurements by Voyagers 1 and 2 and Pioneer 11. The ring structure at 14 radii from Saturn was detected by Vasundhara *et al.*² from the occultation observations of the star SAO158913 on 24 and 25 March 1984. We report here the detection of a ring system at 19 radii using photoelectric observations of the occultation of the star SAO158763 on 12 May 1984.

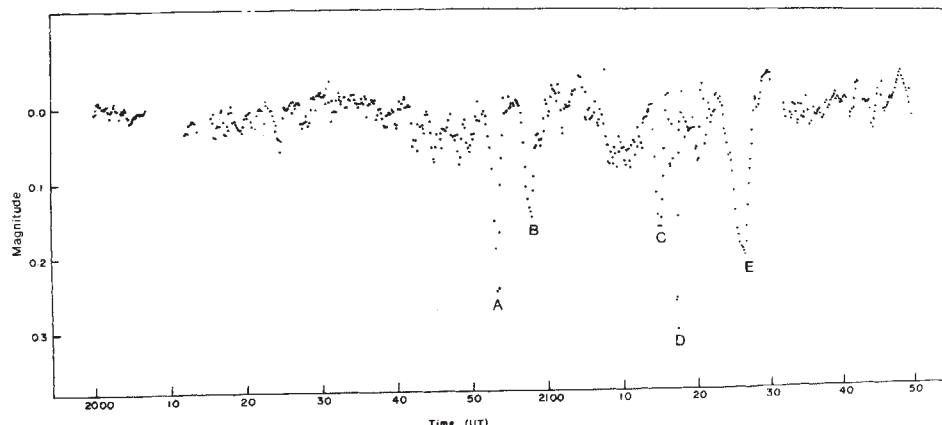
The predicted occultation of SAO158763³ was observed with the 104-cm telescope of the Uttar Pradesh State Observatory (see Fig. 1). The continuous data recorded from 18 h 39 min to 20 h 50 min UT on 12 May 1984 show several prominent dips⁴,

but our observations of predicted occultation of the star on 13 May 1984 from 14 h 27 min to 16 h 30 min UT do not show any such features. The observed variations are within the observational scatter of ± 0.03 mag.

The mean deflections were read from the recorder tracings of 12 May at intervals of 10 s. The light variations shown in Fig. 1 were obtained after correcting for atmospheric extinction. The light diminution is clearly visible between 20 h 38 min and 21 h 29 min UT. The light was almost constant before 20 h 38 min UT and after 21 h 29 min UT to the end of observations at 21 h 50 min UT. Five prominent shallow dips of 0.25, 0.15, 0.16, 0.30 and 0.20 mag were observed at 20 h 53.5 min, 20 h 57.8 min, 20 h 15.2 min, 21 h 17.4 min and 21 h 26.0 min UT respectively, termed A, B, C, D and E zones respectively. The recorded intensities of the star corresponding to these dips are shown in Fig. 2. The most striking features are three prominent spikes at 20 h 53 min 28 s, 21 h 17 min 14 s and 21 h 23 min 48 s UT, of approximate durations of 2.5, 4.0 and 6.5 s respectively.

Assuming that the ring system lies in the equatorial plane of the planet, the distances from the centre of Saturn and the width of the zones A–E and the ring condensations corresponding to spikes 1–3 were calculated and these are given in Tables 1 and

Fig. 1 The occultation light curve observed at Naini Tal of SAO158763 on 12 May 1984. Observations were made using a B filter and EMI 6094S photomultiplier, cooled to -20°C . The photomultiplier output was connected to a photon counting system and data recorded as a continuous series of 0.8-s integrations of photon counts; a strip chart recorder was used to display the signal. The sky transparency was monitored with a photometer attached to the 26-cm finder telescope. A sufficiently wide aperture of about 200 arc s was used to record the brightness of Saturn with the finder telescope. The star was guided with another 20-cm finder telescope. The sky transparency was also monitored photoelectrically by continuously recording the light of the star SAO158808 ($m_v = 6.4$ mag) using a B filter with the 38-cm telescope located about 185 m from the 104-m telescope. After applying corrections for atmospheric extinction, the sky transparency monitored with the two telescopes did not show variations more than ± 0.03 mag throughout the period of observations.



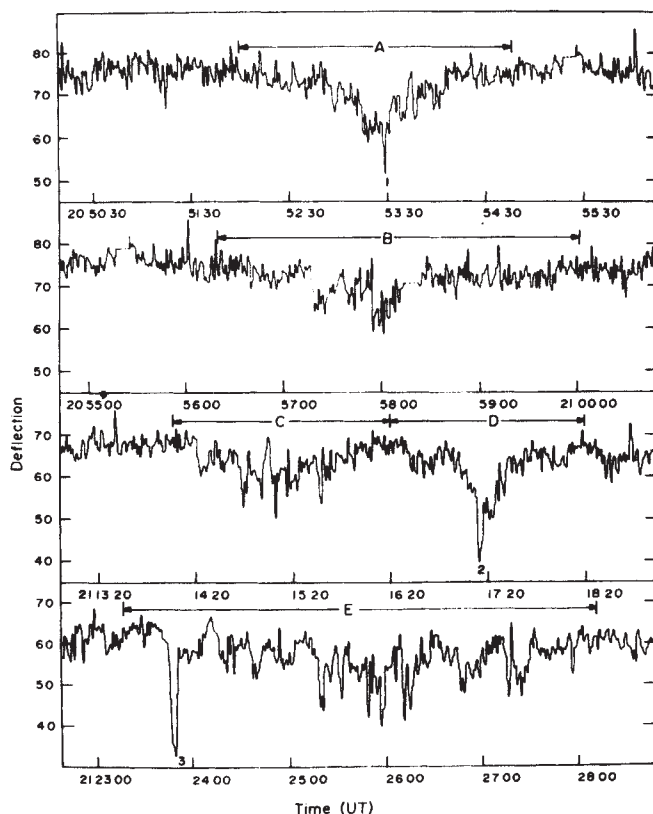


Fig. 2 The observed light variations corresponding to zones A to E.

Table 1 Star occultation events

Ring feature zone	Maximum light diminution (mag)	Mean distance from Saturn (R_s)	Approximate width (km)
A	0.25	20.07	3,100
B	0.15	19.98	4,100
C	0.16	19.66	2,500
D	0.30	19.63	2,200
E	0.20	19.47	5,300

The coordinates and the distance of Saturn were obtained by quadratic interpolation of the coordinates given in *Astronomical Almanac* 1984. The coordinates of the star were taken from the SAO catalogue and coordinates of the star during the event were obtained using the method given in *Astronomical Almanac* (1983, B18).

Table 2 Occultation spike events

Spike no.	Mid-time (UT) h min s	Duration (s)	Maximum light diminution (mag)	Mean distance from Saturn (R_s)	Approximate width (km)
1	20 53 28	2.5	0.48	20.07	45
2	20 17 14	4.0	0.66	19.63	70
3	21 23 48	6.5	0.85	19.51	120

2 for zones and spikes respectively. We obtain a value for the approximate width of the ring structure as 56,500 km, extending from 19.41 to 20.36 R_s . The orbit of Titan is located within the width of this ring structure. The light variations shown in Fig. 1 show that the ring structure can be divided between two broader zones. One broader zone extends from 20 h 38 min to 21 h 04 min UT, with an approximate width of 29,000 km, taking in the two prominent zones A and B; the second extends from 21 h 05 min to 21 h 29 min UT with approximate width 26,400 km, including zones C, D and E.

The shape of the light variation, corresponding to zones A and D, shows that the density of the particulate or gaseous matter changes systematically with a maximum density and a condensation lane near the centre of the zone. The density of the matter in zones B, C and E shows irregular variation, and there are large irregular variations in the distribution of occulting matter in zone E, with signs of the beginnings of a condensation lane.

Received 30 August; accepted 29 October 1984.

1. Lazarus, A. J., Hasegawa, T. & Bagenal, F. *Nature* **302**, 230-232 (1983).
2. Vasundhara, Santhanam & Rozario *IAU Circ.* No. 3941; 3945 (1984).
3. Mink, D. J. *Astr. J.* **88**, 559-561 (1983).
4. Mahra, H. S., Mohan, V., Pandey, A. K. & Sanwal, B. B. *IAU Circ.* No. 3945 (1984).

Turbulent mixing between fluids with different viscosities

I. H. Campbell* & J. S. Turner

Research School of Earth Sciences, Australian National University, GPO Box 4, Canberra 2601, Australia

Fluid dynamic processes associated with the injection of new pulses of magma into magma chambers have received increased attention recently¹⁻³, especially chambers replenished from below with hotter, denser magma of more primitive composition^{3,4}. The results of several laboratory analogue experiments^{3,4} have shown that the dynamical effects of crystallization following replenishment can be quite different when the two layers have similar, or very different, viscosities. The experiments reported below have examined the effect of viscosity differences on the filling process itself. When a turbulent 'fountain' of fluid of low viscosity (ν_1) is injected upwards into a less dense fluid of higher viscosity (ν_2), the two fluids may mix thoroughly or not at all, depending on the relative magnitudes of an input Reynolds number $Re_1 = wd/\nu_1$ and the viscosity ratio ν_2/ν_1 where w is the mean velocity and d is the diameter of the input. The criterion for mixing can be expressed alternatively as $wd/\nu_2 > k$, a constant, in agreement with a more general theoretical argument which is also outlined here. When applied to magma chambers, our results imply that basaltic magmas will mix readily, but that a basaltic inflow will mix with a silicic host magma only with great difficulty.

A very slow input of dense fluid from below mixes little, whatever the viscosity ratio, and forms a distinct layer with a sharp interface above it. This condition has been taken as the starting point for subsequent studies^{3,4} of the cooling and crystallization in the lower layer, followed by overturning. Systematic changes in the petrology and mineral chemistry in sequences of cyclic units in layered intrusions, and the occurrence of chromite seams at the bottom of some of them, suggest, however, that mixing can be important^{5,7}. These observations have led us to consider the 'fountain' model⁸, in which denser fluid is injected upwards as a turbulent jet with excess momentum and falls back in an annular region surrounding the upflow. For aqueous solutions, with comparable viscosities in the fountain and outside it, there is appreciable entrainment, so that the bottom of the tank or chamber becomes stratified. The hybrid layer, which contains a variable mixture of the host and input fluids, has a depth greater than that of the added fluid. The dynamics of these processes can be understood using a combination of previous reversing-plume⁹ and filling-box¹⁰ laboratory models.

Our laboratory experiments show, however, that when a denser fluid (such as K_2CO_3 solution) is injected into a tank of pure glycerol, at an input Reynolds number of about 10^3 , mixing is completely suppressed (see Fig. 1). If the experiment is repeated using water as the host fluid, with the same inflow rate and the same density difference, the two fluids mix thoroughly, as described above. The difference between the two experiments

* Permanent address: J. Tuzo Wilson Research Laboratories, Department of Earth and Planetary Sciences, Erindale Campus, University of Toronto, Mississauga, Ontario, Canada L5L 1C6.

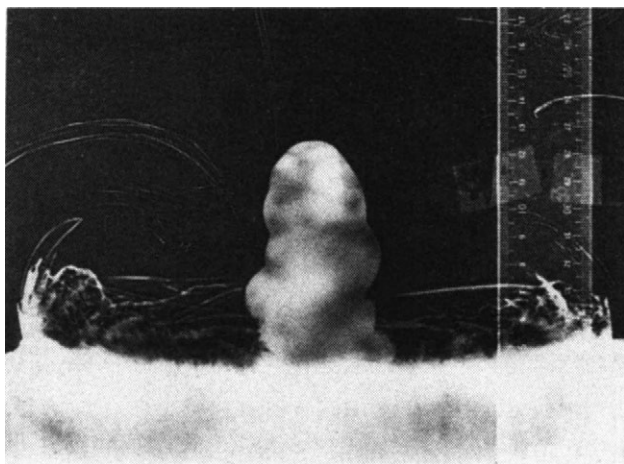


Fig. 1 A fountain of K_2CO_3 solution injected into a tank of less dense glycerol at a value of $wd = 68 \text{ cm}^2 \text{ s}^{-1}$, shown here on a negative print. Note that the fountain is fully turbulent as it passes through the inviscid layer but that its boundaries with the upper viscous layer are smooth. (The crown plume which forms around the fountain results from the double-diffusive transfer of water across the interface.) Scale is in centimetres.

lies in the ratio of the viscosities of the host (ν_2) to input fluids (ν_1) which is 400 in the first case and 0.5 in the second. These observations reveal that ν_2/ν_1 has an important influence on the mixing properties of fountains. But what is the influence of viscosity on mixing in fountains and, given that the inflow remains fully turbulent, what are the criteria for the reduction and final elimination of mixing?

The external parameters, in terms of which the mixing must be described, are the mean velocity w and the diameter d of the input, the kinematic viscosities ν_1 and ν_2 of the input and host fluids, and $g' = g\Delta\rho/\rho$, where $\Delta\rho$ is the density excess of the injected fluid. From these can be formed three independent dimensionless groups; two of these have already been introduced and the third has the form of an internal Froude number:

$$Re_1 = wd/\nu_1; \nu_2/\nu_1; w/(g'd)^{1/2} \quad (1)$$

Alternatively, as developed previously⁹ for the case of equal viscosities, two overall source parameters, the momentum flux $M \propto w^2 d^2$ and the negative buoyancy flux $F \propto g'wd^2$ can be used to form an overall length scale l and a velocity scale v for a turbulent fountain

$$l \propto M^{3/4} F^{-1/2} \quad (2)$$

$$v \propto M^{-1/4} F^{1/2} \quad (3)$$

The total height l_1 of fountains with $\nu_1 = \nu_2$ is proportional⁹ to equation (2), and the constant numerical factor (c_1 say) has been evaluated experimentally (using various values of w , d and $\Delta\rho$ over a much wider range of conditions than in the present experiments).

The product of equations (2) and (3) is independent of F , and proportional to wd , so that Re_1 is an appropriate Reynolds number to describe the whole flow, not only the input conditions. F enters indirectly through its effect on the height of rise. In the experiments described below, both Re_1 and $\Delta\rho$, and hence F , were kept fixed for each series of experiments, and thus the dependence of mixing on ν_2/ν_1 alone was tested directly in a nearly fixed geometry. Although, in principle, the numerical constant c_1 relating l_1 to equation (2) could also depend on ν_2/ν_1 , no such direct viscous effect was detected. The decrease in upward momentum causing the flow in the fountain to reverse was in all cases dominated by buoyancy acting on the inner turbulent flow and the interaction between the upward and downward parts of this inner flow, not by entrainment of the outer fluid.

Consider now the standard description¹¹ of the energy balance and the entrainment mechanism for a turbulent jet entering

surroundings of the same viscosity ν_1 but expressed in terms of the overall 'fountain' parameters (2) and (3). Energy is put into the turbulence at a rate

$$\epsilon \approx v^3/l \quad (4)$$

through large eddies whose length and velocity scales are respectively comparable to the half width and the local mean velocity (corresponding roughly to the large-scale convolutions at the edge of the fountain). External fluid is engulfed by these large eddies, and is then mixed in by smaller scale turbulent motions which are generated by the breakdown of the energy-containing eddies. Energy is transferred by inertial interactions to smaller and smaller scales, and viscosity remains unimportant until the energy is finally dissipated by eddies with length and velocity scales of the order, respectively, $(\nu_1^3/\epsilon)^{1/4}$ and $(\nu_1\epsilon)^{1/4}$. These are the Kolmogorov scales, at which viscosity becomes important: their definition also implies that the Reynolds number based on these dissipating scales is of the order of unity. The rate of entrainment into turbulent fountains at high Reynolds numbers $Re_1 = wd/\nu_1 \propto vl/\nu_1$ is not directly affected by the dissipation scales, that is by viscosity. As Re_1 is reduced below some minimum value, however, the flow ceases to be fully turbulent, and at $Re_1 \approx 0(10)$ the fountain becomes laminar and turbulent entrainment ceases.

When a fluid of high viscosity ν_2 surrounds the turbulent flow and $\nu_2 \gg \nu_1$, the host fluid can be entrained only when the inertial forces associated with the turbulence can distort or wrinkle the bounding surface. The form of the limiting condition, at which viscosity inhibits any such distortion, can be obtained by equating the scale of the energy containing eddies of the turbulent flow to the Kolmogorov scale of the outer fluid which is being entrained by these large eddies:

$$l \approx (\nu_2^3/\epsilon)^{1/4} \quad (5)$$

Eliminating ϵ from equations (4) and (5) gives

$$\frac{vl}{\nu_2} \approx 1 \quad (6)$$

This result can be derived in various (equivalent) ways, for example, by requiring that the normal stresses generated by turbulence be large compared with the viscous stresses generated when the interface is deformed in the manner necessary for mixing. It states directly that viscous and inertial forces are comparable at the scale l in the outer fluid, and it might have been taken as the guiding hypothesis *ab initio*. The derivation through equations (4) and (5) does, however, relate it more firmly to previously accepted ideas about turbulent jets. Using again the result (derived from equations (2) and (3)) that $vl \propto wd$, equation (6) implies that the rate of mixing will be unaffected by any further reduction in ν_2 provided

$$Re_2 = \frac{wd}{\nu_2} > k, \text{ a constant} \quad (7)$$

That is, the criterion for mixing is determined by the input parameters d and w of the fountain and the viscosity ν_2 of the outer fluid, and is independent of ν_1 , the viscosity of the inflow. It can also be written in terms of the inflow Reynolds number as

$$Re_1 = \frac{wd}{\nu_1} > k \frac{\nu_2}{\nu_1} \quad (8)$$

In this form, the relation predicts that as the viscosity ratio increases, the inflow Reynolds number must increase in proportion if mixing is to remain possible.

The above theoretical arguments and particularly the relation (7) provide the basis on which to interpret the experimental results. Three series of experiments have been carried out to test the influence of ν_2 on mixing into fountains, with the other parameters kept constant, as set out in the caption to Fig. 2. For each series, wd and $\Delta\rho$ were kept fixed, the latter being chosen to give an experimentally-convenient height of rise l_1 of the initial fountain. The total amount of fluid supplied to the

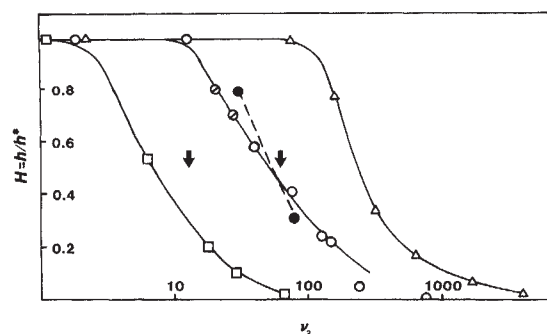


Fig. 2 The results of 'fountain' experiments carried out in a tank $400 \times 196 \times 297$ mm deep, using a fixed nozzle 7.2 mm diameter, and three inflow rates, corresponding to the different symbols. The normalized entrainment $H (= h/h^*)$, as defined in the text) is plotted against the viscosity ν_2 of the host fluid in centistokes, on a logarithmic scale. Δ , $w = 94 \text{ cm s}^{-1}$, $wd = 68 \text{ cm}^2 \text{ s}^{-1}$, $\Delta\rho = 0.10 \text{ g cm}^{-3}$. Input fluid: K_2CO_3 solution, $\nu_1 = 2.46$ centistokes. Host fluid: mixture of glycerol and K_2CO_3 solution. $\circ$, $\bullet$, $w = 25 \text{ cm s}^{-1}$, $wd = 18 \text{ cm}^2 \text{ s}^{-1}$, $\Delta\rho = 0.05 \text{ g cm}^{-3}$. $\circ$, Input fluid: K_2CO_3 solution, $\nu_1 = 2.04$ centistokes. Host fluid: mixture of glycerol and K_2CO_3 solution; $\bullet$, input fluid: NaCl solution, $\nu_1 = 1.06$ centistokes. Host fluid: sodium-carbon-methyl-cellulose (CMC) in water. $\bullet$ (connected by dashed line): fountain height was comparable to experiments with $wd = 68 \text{ cm}^2 \text{ s}^{-1}$. $\square$, $w = 5.4 \text{ cm s}^{-1}$, $wd = 3.9 \text{ cm}^2 \text{ s}^{-1}$, $\Delta\rho = 0.0020 \text{ g cm}^{-3}$. Input fluid: NaCl solution, $\nu_1 = 1.05$ centistokes. Host fluid: CMC in water. The arrows mark the viscosities at which the entrainment is predicted (using equation (8) with $k = 30$) to fall to half its maximum value in the experiments with $wd = 18 \text{ cm}^2 \text{ s}^{-1}$ and $wd = 3.9 \text{ cm}^2 \text{ s}^{-1}$. The displacement of these arrows from the corresponding experimental curves is a measure of the direct effect of ν_1 in reducing entrainment.

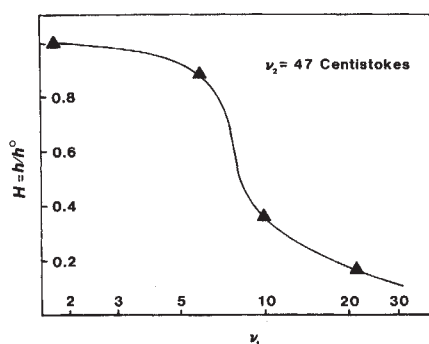


Fig. 3 A plot of the amount of entrainment (H) against ν_1 (in centistokes) on a logarithmic scale, measured in experiments with constant $\nu_2 (= 47 \text{ centistokes})$ and $wd = 18 \text{ cm}^2 \text{ s}^{-1}$. The measured entrainment heights have been normalized to the amount of entrainment (h^0) expected from the previous experiments with $wd = 18 \text{ cm}^2 \text{ s}^{-1}$, as read from Fig. 2.

fountain was also kept constant in each series, and this was recorded as the height h_t of the column of liquid added to the rectangular tank.

The 'entrainment height' h was measured by subtracting h_t from the height h_m of the mixed or hybrid layer at the bottom of the tank, that is $h = h_m - h_t$. (Note that the hybrid layer will now have a gradient of viscosity, as well as density.) The maximum entrainment in each series, equivalent to an entrainment height h^* , was obtained at the minimum value of ν_2 . The results for the different series of experiments (each carried out with different values of l_1 and h_t) can be compared by normalizing h with h^* in each case and defining $H = h/h^*$.

The results of the three series of experiments plotted in this way are summarized in Fig. 2, where $H = h/h^*$ is plotted against ν_2 on a logarithmic scale. Note the following features, which are in accord with the theoretical result given by equation (7):

(1) For any value of wd there is a critical value of $\nu_2 (= \nu_2^c)$

below which changes in the viscosity of the host fluid have no influence on mixing.

(2) Once $\nu_2 > \nu_2^c$ the viscosity of the host fluid has a controlling influence on entrainment, which then falls sharply and becomes very small at values of $\nu_2 > 10\nu_2^c$.

(3) The higher the value of wd , the higher is the value of ν_2 at which the viscosity of the host fluid first influences entrainment.

The slope of these curves (the rate of change of H with $\log \nu_2$), over the range of values of ν_2 for which ν_2 has the strongest influence on entrainment, appears to depend on the geometry of the fountain, in particular the initial fountain height. This height was $l_1 \approx 50$ mm for the experiments with $wd = 3.9 \text{ cm}^2 \text{ s}^{-1}$ and $18 \text{ cm}^2 \text{ s}^{-1}$ and the shape of the curves is also similar for these two cases. In the experiments with $wd = 68 \text{ cm}^2 \text{ s}^{-1}$, however, $l_1 \approx 170$ mm and the slope of the curve is appreciably greater. Two additional experiments were carried out with $wd = 18 \text{ cm}^2 \text{ s}^{-1}$ explicitly to test this dependence on l_1 . The height of the fountain was adjusted (by lowering $\Delta\rho$) to be the same as for the $wd = 68 \text{ cm}^2 \text{ s}^{-1}$ experiments and the slope of the new curve, defined by the solid data points on Fig. 2, does match the higher slope of the $wd = 68 \text{ cm}^2 \text{ s}^{-1}$ relation. These results indicate that the measured value of ν_2^c will depend not only on the considerations leading to equation (7) but also on the geometry of the fountain. Similarly the value of ν_2 at which entrainment falls below some arbitrary fraction, say 10% of h^* , will also depend on l_1 . Note, however, the correspondence between the 'entrainment midpoints' ($H = 0.5$) for the experiments with $wd = 18 \text{ cm}^2 \text{ s}^{-1}$ using the two fountain heights, suggesting that this point may be less sensitive to fountain geometry and its position predictable.

If the relation (7) is valid over the whole range of parameters, a single value of k should describe the $H = 0.5$ points for all the experiments. Taking the experiments with the largest value of $wd = 68 \text{ cm}^2 \text{ s}^{-1}$ as standard, we find $k = 30$. The corresponding viscosities for the other two series of experiments calculated from this value of k are marked with the arrows. The calculated entrainment midpoint for $wd = 18 \text{ cm}^2 \text{ s}^{-1}$ lies at a value of ν_2 20% above the experimentally-determined value, slightly outside the experimental error, but at $wd = 3.9 \text{ cm}^2 \text{ s}^{-1}$ the calculated viscosity is almost twice that of the experimental value.

The substantial agreement at the higher wd values supports the prediction contained in equation (7), that entrainment at high inflow Reynolds numbers is controlled by ν_2 and is independent of ν_1 . We emphasize again that equation (7) is valid only when the input fluid is fully turbulent and $\nu_2 \gg \nu_1$, and that the effect of ν_2 is indirect, acting through its influence on the distortion of the interface. The increasing disagreement at lower wd values is explained by the fact that Re_1 is then so low that the turbulence in the fountain is being directly damped by ν_1 . ν_2 and ν_1 are also more nearly comparable over the critical range of values where ν_2 influences entrainment. This hypothesis was tested by carrying out a series of experiments in which ν_1 was varied, at a constant value of ν_2 , as plotted in Fig. 3. The results show that ν_1 had a minor influence on entrainment until ν_1 reached a value such that $Re_1 \approx 200$ and $\nu_2/\nu_1 \approx 8$. At values of ν_2/ν_1 below 8 there was a sharp decrease in entrainment, and ν_1 then had a dominant influence on the entrainment process. Experiments at higher values of wd and ν_2/ν_1 than the maximum reported here (which are not easily achieved in our laboratory) should eliminate these direct effects of ν_1 and produce better agreement between experiment and the high Reynolds number theory.

The main results of this work can be summarized thus: when $Re_1 \geq 400$, ν_1 has no significant effect on mixing, which is controlled by ν_2 . In this regime, the value of $Re_2 = wd/\nu_2 = k$ at which the mixing is reduced by 10% relative to its value at large Re_2 is $k \approx 70$ and a 50% reduction is attained at $k = 30$.

The experiments provide a plausible first test of the theoretical ideas presented and give us some confidence that the principles discussed will be more generally applicable. Equations of the form (7) or (8) should apply also to turbulent jets and plumes passing through a more viscous host fluid, although the numeri-

cal factors k will probably vary with the type of flow considered. When the criterion for mixing, $wd > kv_2$, is applied⁸ to natural magmas, the conclusion is very clear. Our results show that if a primitive basaltic magma is injected into a chamber containing fractionated basaltic magma, the two magmas will mix readily. However, if basaltic magma is injected into a chamber containing a much more viscous silicic melt, little or no mixing will occur.

We thank Professor G. K. Batchelor, Dr H. E. Huppert and Dr G. N. Ivey for discussions which helped to clarify our theoretical ideas.

Received 18 July; accepted 24 October 1984.

1. Sparks, R. S. J., Meyer, P. & Sigurdsson, H. *Earth planet. Sci. Lett.* **46**, 419–430 (1980).
2. Huppert, H. E. & Sparks, R. S. J. *Contr. Miner. Petrol.* **75**, 279–289 (1980).
3. Huppert, H. E. & Turner, J. S. *Earth planet. Sci. Lett.* **54**, 144–152 (1981).
4. Huppert, H. E., Sparks, R. S. J. & Turner, J. S. *Earth planet. Sci. Lett.* **57**, 345–357 (1982).
5. Huppert, H. E., Sparks, R. S. J. & Turner, J. S. *Earth planet. Sci. Lett.* **65**, 377–381 (1983).
6. Campbell, I. H. *J. Petrol.* **18**, 183–215 (1977).
7. Campbell, I. H. & Murck, B. *J. Petrol.* (in the press).
8. Campbell, I. H. & Turner, J. S. *J. Petrol.* (in the press).
9. Turner, J. S. *J. Fluid Mech.* **26**, 779–792 (1966).
10. Baines, W. D. & Turner, J. S. *J. Fluid Mech.* **37**, 51–80 (1969).
11. Batchelor, G. K. *The Theory of Homogeneous Turbulence* (Cambridge University Press, 1953).

Discovery of permanent Amazon lakes and hydraulic disturbance in the upper Amazon Basin

P. A. Colinvaux*, M. C. Miller†, Kam-biu Liu*,
M. Steinitz-Kannan‡ & I. Frost*

* The Ohio State University, Department of Zoology, Columbus, Ohio 43210, USA

† University of Cincinnati, Department of Biological Sciences, Cincinnati, Ohio 45221, USA

‡ Northern Kentucky University, Department of Biology, Highland Heights, Kentucky 41076, USA

We describe here the first autonomous permanent deep lakes found in the Amazon lowlands. Permanent Amazon lakes have hitherto been considered unlikely because the ubiquitous rivers meandering over flat terrain should obliterate ancient basins, resulting in lakes subject to annual flooding, the so-called 'varzeas'^{1–4}. The lakes described here are in the rain forests of Ecuador in western Amazonia, within 100 km of the Andes mountains, but at elevations below 700 m (Fig. 1). They occupy either deep abandoned channels of high-energy rivers or closed volcanic basins. Sedimentary histories and pollen data over several millennia from these lakes record disturbance of the Amazonian ecosystem by flooding, encouraging the view that high Amazonian species-richness is maintained by non-equilibrium processes^{5,6}.

The lakes were found in field surveys in the summers of 1982 and 1983. We used motorized rubber boats to make bathymetric maps with a recording sonar, to sample for plankton, water chemistry, ¹⁴C productivity, thermal stratification and light penetration and to raise sediment cores with a Livingstone piston sampler. Tables 1–4 give radiocarbon ages, water chemistry and physical structure, productivity and common algae for both sets of lakes.

Lakes Ayauchi¹ and Kumpak^a are in remote parts of southern Ecuador, at the foot of the eastern cordillera of the Andes, between the 300 and 700-m contours. We give the names used by the Shuara (Jivaro) people native to the region. Both basins are closed, encircled by rims and have concentric bathymetry.

The rim of Ayauchi¹ is steep and includes embedded boulders which appear volcanic. A sediment core reveals a base layer of feldspar-rich pyroclastic deposits with pebbles of pumice underlying 3.6 m of black organic gyttja. A radiocarbon age of $7,010 \pm 130$ yr at the base of the gyttja is the oldest date yet obtained from the Amazonian lowlands (Table 1).

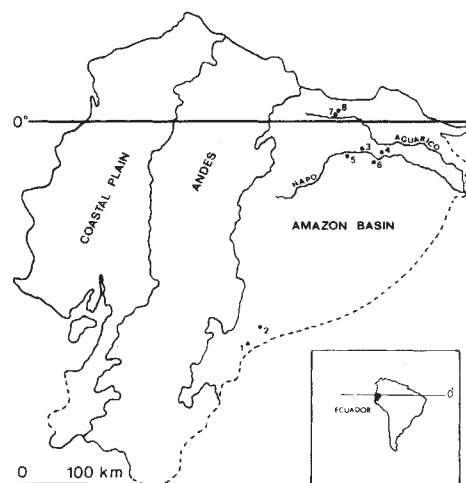


Fig. 1 Location of lakes in Ecuador. 1, Ayauchi¹; 2, Kumpak^a; 3, Limoncocha; 4, Garzacocha; 5, Taracoa; 6, Añangucocha; 7, Santa Cecilia; 8, Lago Agrio.

Table 1 Radiocarbon dates from Amazonian lake sediments

Lake	Core	Depth (cm)	¹⁴ C date (yr BP)
Ayauchi ¹	(A)	352.0–362.0	7,010 ± 130
Kumpak ^a	(A)	461.0–476.0	1,900 ± 140
Kumpak ^a	(A)	952.0–965.0	4,090 ± 100
Kumpak ^a	(B)	968.0–982.0	3,840 ± 220
Kumpak ^a	(B)	1,264.0–1,276.0	4,040 ± 130
Kumpak ^a	(B)	1,846.0–1,862.0	5,120 ± 100
Limoncocha	(A)	160.0–173.0	755 ± 135
Limoncocha	(B)	169.5–180.0	960 ± 140
Limoncocha	(B)	555.0–565.0	1,230 ± 120
Añangucocha		112.5–118.0	830 ± 160
Añangucocha		279.0–283.0	1,320 ± 160
Añangucocha		315.5–319.0	2,170 ± 150
Añangucocha		344.5–348.5	3,140 ± 140
Añangucocha		444.5–449.5	3,020 ± 130
Añangucocha		484.0–494.0	3,040 ± 80
Santa Cecilia		331.0–342.0	610 ± 100
Santa Cecilia		552.0–564.0	785 ± 100
Lago Agrio	(D)	126.0–135.0	675 ± 190

The Kumpak^a sediments revealed by an 18.6-m core under 19.5 m of water are allochthonous clay with organic matter. The core is banded throughout with alternating dark and pale layers of clay-rich gyttja interspersed with many distinctive clay laminae several millimetres to 12 cm thick, together with a few thin bands of tephra. X-ray diffraction of the clay laminae shows sharp reflectivities at 20.8θ and 26.66θ (both quartz), but with little or no clay minerals. This suggests that the Kumpak^a catchment is made of young quartz-rich volcanic material. Both lakes can be seen from the air to lie significantly above the neighbouring rivers. We conclude that both occupy explosion craters (maars), making them the first volcanic lakes identified in the Amazon Basin.

Ayauchi¹ appears blue from the air, although overhung by rain forest on all sides. It has a secchi disk transparency of 4.5 m, oxygen to 20 m, and a maximum depth of 40 m (Table 2). Thermal stratification is minimal, the temperature falling only 1°C in 25 m. ¹⁴C productivity of $84\text{ mg C m}^{-2}\text{ h}^{-1}$ is not abnormally low for Amazonian lakes, although allowing the designation 'oligotrophic' (Table 3). Algal biomass, however, is low. This, coupled with an absence of humic colour and suspended mineral particles, allows light to penetrate, giving a blue colour to the water, unique for a rain forest lake. We suggest that the deep penetration of oxygen is a function of the consequent photosynthesis in the deeper parts of the lake.

Table 2 Physical and chemical parameters of permanent lakes of the upper Amazon

Parameter	Lake								
	Ayauchi	Kumpak ^a 1982	Kumpak ^a 1983	Limoncocha	Garzacocha	Taraoa	Añangucocha	St Cecilia	Lago Agrio
Latitude	2°5' S	3°2' S	3°2' S	0°24' S	0°35' S	0°33' S	0°40' S	0°4' N	0°7' N
Longitude	78°1' W	77°49' W	77°49' W	76°38' W	76°20' W	76°45' W	76°25' W	77°1' W	76°55' W
Area (km ²)	0.04	1.29	1.29	2.04	0.28	0.40	0.16	0.06	0.27
Max depth (m)	40.00	19.00	19.00	3.10	2.30	2.20	4.30	7.50	3.50
Conductivity (μΩ cm ⁻¹)	283.00	60.00	110.70	118.00	75.00	373.00	49.00	70.00	69.00
Alkalinity (mg l ⁻¹)	0.24	0.62	0.57	0.99	0.40	0.37	0.24	0.67	0.68
pH	6.47	6.80	7.20	9.00	7.40	5.90	6.00	7.60	6.30
Surface temperature (°C)	28.20	26.50	26.90	28.00	28.00	26.50	26.80	26.00	28.70
Bottom temperature (°C)	27.50	23.50	23.00	25.00	25.00	23.80	22.90	24.70	24.80
Secchi depth (m)	4.50	0.98	1.85	0.51	0.75	0.80	1.25	2.00	1.07
NH ₄ ⁺ (mg N l ⁻¹)	1.00	0.93	2.46	0.33	0.87	0.86	0	1.72	1.38
NO ₃ ⁻ (μg N l ⁻¹)	11.60	—	0.80	—	—	—	—	—	—
PO ₄ ³⁻ (μg P l ⁻¹)	3.73	36.60	10.81	75.00	138.00	108.00	60.00	69.00	113.00
SO ₄ ²⁻ (mg S l ⁻¹)	—	2.30	—	4.30	0.06	0.07	0.03	6.45	6.30
Si ⁴⁺ (mg l ⁻¹)	3.50	1.10	3.40	31.50	21.10	4.10	5.50	24.80	24.60
Cl ⁻ (mg l ⁻¹)	5.57	—	5.35	—	—	—	—	—	—
Ca ²⁺ (mg l ⁻¹)	0.97	7.81	10.90	4.20	2.42	8.26	1.10	1.99	1.26
Mg ²⁺ (mg l ⁻¹)	1.23	0.36	0.55	4.98	2.31	3.22	0.01	2.36	2.35
K ⁺ (mg l ⁻¹)	0.83	1.17	0.33	2.17	1.48	2.07	0.61	1.35	1.61
Na ⁺ (mg l ⁻¹)	1.80	1.22	1.00	6.43	4.05	48.70	3.50	4.48	8.46
Fe ³⁺ (mg l ⁻¹)	0	1.40	0.23	0.65	2.32	1.47	0.34	0.27	2.96
Al ³⁺ (μg l ⁻¹)	0	1.02	0	31.46	21.10	4.15	5.50	24.90	24.55
Zn ²⁺ (μg l ⁻¹)	55.70	15.35	41.20	79.90	25.60	14.80	12.50	11.40	21.30

Table 3 Productivity measurements of selected upper Amazon lakes

Parameter	Lake								
	Ayauchi	Kumpak ^a 1982	Kumpak ^a 1983	Limoncocha	Garzacocha	Taraoa	Añangucocha	St Cecilia	Lago Agrio
a, Average chlorophyll <i>a</i> concentration in photic zone (mg m ⁻³)	2.09	10.27	7.21	5.89	—	9.09	5.03	28.20	10.30
Primary production: b, mg C m ⁻² h ⁻¹	84.26	61.80	220.10	137.00	35.10	157.10	163.60	193.40	103.50
c, Estimated g C m ⁻² day ⁻¹	1.37	0.48	2.19	1.62	0.42	1.35	1.55	2.54	1.36
Ratio b/a (production chlorophyll ⁻¹ h ⁻¹)	40.33	6.02	30.53	23.36	—	17.28	32.52	6.86	10.05

Water chemistry of Ayauchi suggests that available nutrients are not significantly different from those in other Amazonian lakes (Table 2). The algae, however, are different, as this is our only desmid lake (*Spinocostarium* sp. 60% total algae, Table 4). A high proportion of desmids is typical of soft-water oligotrophic lakes elsewhere⁷, although the water is usually acidic and coloured with humus, unlike Ayauchi. Our working hypothesis is that the special properties of Ayauchi result from the complete absence of mineral inputs from the watershed. The lake has no inlets and the sediment core confirms that no allochthonous minerals enter the lake. As a result, the water is never turbid nor charged with allochthonous nutrients at the surface. Light penetrates, desmids are favoured, biomass is minimal and classical oligotrophic properties result.

Kumpak^a is a turbid-water lake with narrow inlet streams entering on all sides. It is larger than Ayauchi (Table 2) and drains a much larger catchment. Radiocarbon dating of our cores shows that sediment collects rapidly at rates up to 1 cm yr⁻¹ (Table 1). Following one heavy rain storm, we witnessed clouds of muddy water entering the lake from inlet streams; this suggests that the textural bands represent stormy episodes that inject allochthonous material and nutrients into the surface water.

In June–July of both 1982 and 1983, temperature stratification in Kumpak^a was as slight as in Ayauchi, but secchi disk transparency was only 1.85 m and water was anoxic below 1 m (Table 2). Productivity was comparable with that of riverine lakes, although one measurement following a rain storm was the highest in our suite of lakes. The production-to-biomass ratio was low, as in the riverine lakes (Table 3), and the phytoplankton

was comparable with the most productive of them, Limoncocha (Table 4). Secondary production also seems high, as the oxygenated surface water supported dense (always visible) populations of small (2–10 cm) fish, especially small piranhas. Caimans were plentiful, unlike in Ayauchi where we saw none.

We conclude that the presence or absence of allochthonous mineral inputs determines the trophic state of these Amazonian crater lakes. Rapid sedimentation in Kumpak^a not only injects nutrients but also produces extensive shallows where nutrients can be regenerated. The blue-water properties of Ayauchi result because the fetch across the catchment is too small for run-off water to penetrate the forest litter, thus denying the lake allochthonous minerals. Both lakes are warm, discontinuous polymictic lakes; thus circulation occurs infrequently and nutrients are trapped in the deep water.

The Ecuadorian lakes along the Napo and Aguarico rivers differ from these lakes of central Amazonia in that they are deeper, particularly in the less pronounced dry season, and occupy steep-sided channels cut by high-energy rivers (Fig. 1, Table 2). Some have been in existence for at least 3,000 radiocarbon years (Table 1).

All our lakes are black water except Limoncocha. Daily or seasonal patterns of mixis are probably not markedly different from the *varzea* lakes⁸, although the greater depths of our lakes in the dry season suggest more prolonged stratification and perhaps multiple thermoclines within the epilimnion⁹. Episodes of anoxic surface water following complete overturn, as happens in central Amazonia, are unlikely.

Daily productivities were estimated from short-term incubations corrected by the proportion of the sine curve of daily

Table 4 Most common algae in permanent lakes of the upper Amazon

	Añangucocha	Ayauchi	Garzacocha	Kumpak ^a	Lago Agrio	Limoncocha	Santa Cecilia	Taracoa
Cyanophyta								
<i>Anabaena</i>	+++					+	+++	+++
<i>Aphanocapsa</i>						+	+	
<i>Chroococcus</i>	+							
<i>Merismopedia</i>					+		+	
<i>Microcystis</i>		+	+	+++		+++		
<i>Oscillatoria</i>	+					+	+	
Chlorophyta								
<i>Arthrodesmus</i>		+						
<i>Binuclearia</i>	+							
<i>Botryococcus</i>		+		+				
<i>Characium</i>						+		
<i>Chlamydomonas</i>			++		+			
<i>Chlorella</i>						+		
<i>Cosmarium</i>	+							
<i>Desmidium</i>	+							
<i>Euastrum</i>	+							
<i>Eudorina</i>	++	+						
<i>Gonatozygon</i>	+							++
<i>Micrasterias</i>	++						+	
<i>Mougeotia</i>	+							
<i>Oedogonium</i>	+						+	
<i>Ophiocytium</i>					+			
<i>Pandorina</i>						+		
<i>Spinocosmarium</i>		+++						
<i>Spirogyra</i>	+							
<i>Staurastrum</i>		+					+	
<i>Tetrahedron</i>		+					+	
<i>Westella</i>					+			
Bacillariophyta								
<i>Melosira granulata</i>			+++		+++			
<i>Tabellaria</i>				++			+	
<i>Pennates</i> (other)	+		++			+	+	
<i>Centrics</i> (other)			+	+				
Euglenophyta	+							+
Pyrrophyta								
<i>Peridinium</i>				+				

†, 1–20%; ++, 20–50%; +++, 50–99%.

illumination during which the incubation took place (Table 3). Productivity was positively related to ammonia concentration and negatively to soluble reactive phosphorus, water colour and soluble iron. None of these relationships was significant statistically, but trends were obvious in our sample of eight lakes. Water colour inhibits light penetration in the black-water lakes, reducing photic depth. The total iron is bound or peptidized by the coloured humic materials and hence is inversely related to light and production. These data suggest that productivity is simply light-limited, as is conventionally suggested for black-water systems. However, more is involved than simple occlusion of light, because secchi disk transparency is not closely correlated to productivity.

In both the riverine and crater lakes, the prime determinants of productivity are allochthonous. When particulate allochthonous inputs are almost totally absent, as in Ayauchi, a system of deeply scattered algae, predominantly desmids, achieves a high production-to-biomass ratio. Heavy inputs of allochthonous humic complexes in black-water riverine lakes imply special limits and are associated with high densities of Cyanophyta or *Melosira granulata* (Table 4). Turbid inputs of mineral sediments allow greater productivity (Kumpak^a, Limoncocha) but are associated with light occlusion and, perhaps, with rapid sequestering of nutrients in sediments in the deep water. Production is modest in all three kinds of tropical systems, however, probably because the soluble nutrients in watersheds are low, as has long been argued for the lowland wet tropics.

Cores from four riverine lakes (Lago Santa Cecilia, Lago Agrio, Añangucocha, Limoncocha, Fig. 1) show that they have deposited gyttja without riverine sediments for the past 800 radiocarbon yr and have therefore been isolated from the rivers during this time. Under this gyttja, however, all four lakes have a deposit of clay and fine sand similar to sediments of the modern rivers. Radiocarbon ages of about 1,300 yr were obtained from the bottom of this unit in two of the lakes (Table 1). In Añangucocha, the riverine sediments are underlain by a thick

organic deposit spanning to 3,000 yr ago. These data strongly suggest that all four lakes were flooded by the parent rivers between 1,300 and 800 yr BP but have been isolated ever since.

Pollen data from Añangucocha are consistent with the stratigraphical evidence for flooding, suggesting inundation of old swamp sites and the simultaneous creation of riparian habitats along extended river channels¹⁰. Absy¹¹ postulates the existence of a relatively wet period from 2,000 to ~1,100 yr ago for central Amazonia, followed by a drier phase culminating around 700 yr ago. Our flooding episode seems to be slightly out of phase with Absy's wet period, but her time control comes from only one ¹⁴C date and remains to be substantiated. Other studies also suggest widespread flooding in the upper Amazon before 800 yr ago^{12–14}. Our radiocarbon chronology for the four lakes spread on a west-east axis of ~100 km appears to define this event as spanning between 1,300 and 800 yr ago. The most likely cause of this flooding is increased precipitation on the eastern slopes of the Andes, the effects of which would have been felt throughout the Amazon drainage system.

The 'intermediate disturbance hypothesis'¹⁵ suggests that the high species diversity of tropical rain forests is maintained because moderate disturbance prevents competitive exclusion but is insufficient of itself to sustain a high exclusion rate. The proxy manifestation of forest disturbance is recurrent tree fall that makes forest life a process of continual succession in gaps^{5,6}. Large-scale floods such as we have documented, at intervals of <1,000 yr, should prevent competitive exclusion of forest trees, at least in the areas directly affected, thus being an important source of disturbance leading to high diversity.

Forests of the Amazon Basin may be more disturbed than other tropical rain forests because of their unusual hydraulic stress. Even the so-called *terra firma* areas above the present river flood plains are essentially the higher ridges of a continuously-eroded surface and themselves subject to denudation, as shown by the massive erosion of the Kumpak^a catchment. If diversity is proportional to disturbance over some intermediate

range, then the long history of hydraulic disturbance in Amazonia may have been a prime contributor to the well known high diversity there. This hypothesis predicts that tropical rain forests that are not on ancient peneplains in steady state should be less diverse than Amazonia.

We thank H. and T. Steinitz, R. Vargas, J. Vela, and C. Segundo, who provided our bases in Ecuador, NSF for funding, and the Ruales family of pilots, M. Espinsa, K. Olson, S. Longenbaker, M. Scheutzwow and C. Vasquez for assistance.

Received 11 July; accepted 8 October 1984.

1. Marlier, G. *Amazonia* 1-2, 91-115 (1973).
2. Irion, G. *Trop. Ecol.* 23, 76-85 (1982).
3. Sioli, H. *Atlas Simp. Sobre a Biota Amazonica* 3, 9-50 (1967).
4. Hill, G. & Rai, H. *Arch. Hydrobiol.* 96, 99-105 (1982).
5. Connell, J. H. *Science* 199, 1302-1310 (1978).
6. Hubbel, S. P. *Science* 203, 1299-1308 (1979).
7. Hutchinson, G. E. *Treatise on Limnology* Vol. 1 (Wiley, New York, 1957).
8. Tundisi, J. G. et al. *Hydrobiologia* 108, 3-15 (1984).
9. Lewis, W. M. *Zooplankton Community Analysis* (Springer, Berlin, 1979).
10. Frost, I. thesis, Ohio State Univ. (1984).
11. Absy, M. thesis, Univ. Amsterdam (1979).
12. Meggers, B. J. *Ecuador* (Praeger, New York, 1966).
13. Van der Hammen, T. & Gonzalez, E. *Leid. geol. Meded.* 32, 183-191 (1965).
14. Wijmstra, T. A. *Leid. geol. Meded.* 39, 261-267 (1967).
15. Connell, J. H. *Science* 204, 1344-1345 (1979).

Hadar AL 162-28 endocranial as evidence that brain enlargement preceded cortical reorganization in hominid evolution

Dean Falk

Department of Anatomy/Caribbean Primate Research Center,
University of Puerto Rico School of Medicine, PO Box 5067,
San Juan, Puerto Rico 00936

On the basis of a description¹ of an endocranial from Hadar early hominid AL 162-28, it has been suggested that cerebral organization towards a human pattern occurred as early as 3-4 Myr ago. I have studied a cast of the AL 162-28 calvaria and a copy of an endocranial prepared from the original fossil, and report here observations regarding cranial capacity, the relationship between endocranial and skull, sulcal pattern, brain shape and cranial venous sinuses. Contrary to the earlier report¹, all of these features appear to be consistent with an ape-like external cortical morphology in Hadar early hominids and in my view there is no evidence for expansion or reorganization of parietal/occipital regions. Cranial capacity of AL 162-28 is at least 10% and 29% smaller than respective mean capacities of subsequently living gracile and robust australopithecines, who also exhibit ape-like cortical patterns. Thus palaeoneurological evidence from the entire early hominid record suggests that the trend towards brain enlargement preceded cortical reorganization.

Separate hemicasts were prepared from the left (most extensive) hemisphere of the AL 162-28 endocranial and the right hemisphere of the Taung endocranial, whose cranial capacity (uncorrected for age) is 404 cm³ (refs 1, 2). A right hemicast was also prepared from an orangutan endocranial of 338 cm³ cranial capacity. The Hadar hemicast was then compared with the complementary hemicasts from the Taung and *Pongo* specimens and its capacity found to fall between the two volumes. Because the AL 162-28 specimen lacks both frontal and temporal regions, the shape of the entire hemicast cannot be reconstructed with certainty. However, the available occipital/cerebellar region is shaped more like the corresponding area in the *Pongo* endocranial than that in the Taung specimen (see below). Holloway estimated the capacity of AL 162-28 at "just less than 400 cm³"¹. I agree that the capacity is under 400 cm³, but prefer a less precise estimate of 350-400 cm³. However, it is important to note that even if the capacity of AL 162-28 were as large as 400 cm³, this would still be the smallest adult cranial capacity in the hominid fossil record to date.

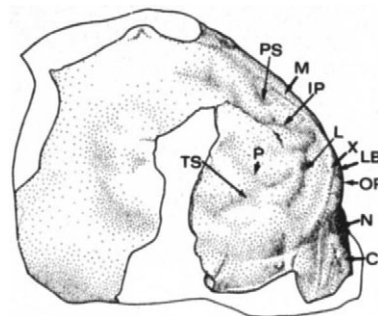


Fig. 1 Left lateral view of endocranial from Hadar hominid AL 162-28. Note that the cerebellum has a more caudal projection than does the occipital pole. Since traces of the lambdoid suture in the cranium correspond with ridge X laterally on the endocranial, and are caudal to it medially, the lambdoid suture is located caudal to L. See text for discussion. Abbreviations: C, cerebellum; IP, intraparietal sulcus; L, lunate sulcus; X, ridge reproduced under nuchal crest; LB, location of trace of lambdoid suture caudal to the medial end of X; M, endocranial midline; N, non-bony matrix (dark stippling); OP, occipital pole; PS, superior parietal sulcus; TS, superior temporal sulcus; P, superior parallel branch of TS.

Cranial capacities have recently been published for two additional Hadar early hominids³. A capacity of 500 cm³ was given for partial adult cranium AL 333-45, but this estimate was determined from an endocranial whose entire frontal portion (which varies in shape within hominoids) was reconstructed in plasticine³. AL 333-105, on the other hand, is a partial juvenile cranium⁴ that appears dentally to be at most 2 yr younger than the Taung specimen^{4,5}. Although the two latter specimens appear to be nearly the same age, Taung's cranial capacity of 404 cm³ is 26% larger than the estimated capacity of 320 cm³ for AL 333-105³. If AL 333-105 were about 5 yr old, as suggested by its dentition^{4,5}, its cranial capacity would be approximately 90% of its likely adult value⁶, which would therefore be 352 cm³. (This estimate may actually be too large because it is based on human data—a 5-yr-old chimpanzee would already have obtained its adult cranial capacity⁶.)

The mean cranial capacity of adult gracile australopithecines is estimated² to be 442 cm³, while the value for robust australopithecines⁷ is 516 cm³. Thus the minimum increase in cranial capacity between AL 162-28 and the average adult gracile australopithecine is 10.5%; the minimum increase between AL 162-28 and the mean adult robust australopithecine is 29%. (If the capacity of AL 162-28 were as low as 350 cm³, the increases would be 26% and 47% respectively.) In sum, the small cranial capacities of two of the three Hadar specimens for whom estimates are available (AL 333-105 and AL 162-28) suggest that discussions of stasis in brain enlargement before the South African australopithecines, or statements that brain size was the last element to change in human mosaic evolution¹, are not supported by the available evidence.

The AL 162-28 endocranial was prepared from a partial calvaria composed of portions of both parietals and the major portion of the occipital squama⁴. My observations of a cast of the fossil confirm the published observations⁴ that "the external aspect of the lambdoid suture is nearly completely obliterated," ... "the internal aspect of the lambdoid suture exhibits advanced fusion" ... and "the internal aspect of the sagittal suture is completely obliterated." Thus, AL 162-28 represents an adult specimen with well-knit sagittal and lambdoid sutures.

A dramatic ridge, the medial end of which is directed rostral to a small remnant of lambdoid suture, crosses the left hemisphere of the endocranial in a medial-lateral direction (X in Fig. 1). Comparison of the endocranial with the cast of the skull reveals that this ridge is reproduced from the internal surface of the calvaria directly under the nuchal crest (including its left lateral third which corresponds to the compound temporonuchal crest and a small portion of its medial continuation into the temporal line⁴). Similar ridges appear on one other hominid endocranial,

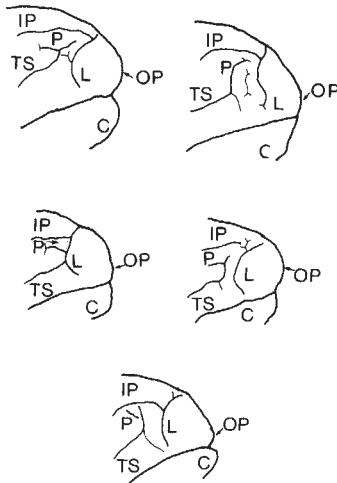


Fig. 2 Left lateral views of chimpanzee brains modified after Connolly¹⁴. Abbreviations as in Fig. 1. Compare with Fig. 1 and note relationship between IP and L.

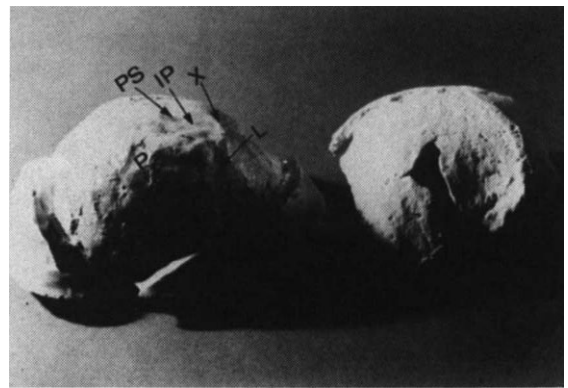


Fig. 3 A copy of the AL 162-28 endocast (left) next to a copy of the calvaria from which the endocast was produced (right). Abbreviations as in Fig. 1. Note that the ridge (X) on the endocast corresponds to the nuchal crest on the calvaria, and that figure should be rotated so that labels are vertical in order for orientation of specimens to be comparable to that of Fig. 1.

KNM-ER 1805 from Kenya^{8,9}. Comparison of this endocast with the original specimen¹⁰ reveals that the right and left ridges were reproduced from the internal surface of the calvaria directly under the parasagittal crests and their compound temporonuchal extensions.

Several sulci are reproduced on the left hemisphere of the AL 162-28 endocast (Fig. 1). Holloway's identification¹ of the sulcus that courses posterior-medially in the parietal region as the intraparietal sulcus (IP) is confirmed. However, as IP courses rostrally, it merges with a crack, and its rostral termination (and therefore its length) is uncertain. A depression (PS) medial to IP may represent the superior parietal sulcus. The superior temporal sulcus (TS), including its superior parallel branch (P), is also reproduced on the endocast.

The caudal end of IP terminates at another sulcus that courses parallel and approximately 5 mm rostral to ridge X. This sulcus is identified as the lunate sulcus (L) for the following reasons. First, its groove-like appearance is typical of sulci that are reproduced on endocasts. Second, in keeping with allometrically based expectations¹¹⁻¹³, the orientation and rostral location (see below) of this sulcus correspond to the orientation and rostral location of lunate sulci in other similarly sized primate brains^{13,14}. Third, the configuration of IP merging caudally with L is typical of similarly sized pongid brains¹⁴ (see Fig. 2).

The feature that I believe to be the lunate sulcus has previously been called "groove B", and stated to be "a depression caused by the inferior lip of the posterior portion of the parietal bone at the lambdoid suture"¹. The possibility of the feature in question being the lunate sulcus was not discussed, and it was stated that "there are no landmarks at the posterior end of groove A (IP) which can be interpreted as a pongid-like lunate sulcus"¹. I have carefully examined the cast of the AL 162-28 calvaria to see if the feature that I have identified as L might not instead be an artefact due to lipping at the lambdoid suture (Fig. 3). I failed to detect any lipping, and my observations instead confirm those of Kimbel *et al.*⁴ that "the internal aspect of the lambdoid suture exhibits advanced fusion" ... and that "very faint sutural traces are intermittently visible on the left" in this adult specimen with highly obliterated sutures⁴.

The lambdoid suture is also located too far caudally relative to L for the latter to be explained as an artefact due to lipping at this suture. The internal surface of the calvaria reveals traces of the lambdoid suture directly underneath the nuchal crest, which corresponds with Kimbel *et al.*'s observation that "near asterion the [external aspect of the] suture is superimposed on the lateral extremity of the nuchal crest"⁴. As noted above, the nuchal crest is directly over all but the most medial aspect of ridge X (to which it and a trace of the lambdoid suture are both caudal). Because ridge X is 5 mm caudal to L, one may surmise

from traces of the lambdoid suture on both the endocast and calvaria that it courses at least 5 mm caudal to L throughout its extent. The relationship between L and the lambdoid suture is of some importance, because that seen in specimen AL 162-28 is typical of pongid brains but not of human brains¹⁴. In humans, at least the lateral portion of L (if not the entire sulcus) is located caudal to the lambdoid suture¹⁴.

Figure 1 reveals an interesting feature of brain shape in AL 162-28: the cerebellum projects further posteriorly than does the occipital pole. This relationship is consistent with the fact that "the cerebellar fossae appear to be deeper than the cerebral fossae"⁴, and may not be attributed to distortion of the fossil because AL 162-28 "exhibits virtually no distortion"⁴. In humans, the occipital cortex is expanded over the cerebellum so that the occipital pole is the most posterior point visible in lateral views of the brain¹⁴. In pongids, in whom the parietal/occipital cortex is less caudally expanded than in man, the most posterior point may be either the occipital pole or a point on the cerebellum¹⁴ (Fig. 2). The shape of the caudal portion of the AL 162-28 endocast is therefore inconsistent with a human-like shape, and this feature, in combination with the rostral position of the lunate sulcus, contradicts the conclusion that AL 162-28 exhibits "expansion of parietal association cortex and the diminution of lateral extent of primary visual striate cortex ..."¹.

I agree with Holloway's observations¹ that "the transverse sinus cannot be discerned between the occipital and cerebellar lobes, nor is the fragment complete enough to detect any marginal or accessory sinus ...". I would further add that the apparent lack of transverse sinuses suggests that if the foramen magnum region were preserved in AL 162-28, an enlarged accessory sinus system would be apparent¹⁵⁻¹⁷.

Taken together, the data suggest that the endocast of AL 162-28 is small, simple and ape-like, and the following combination of features reproduced on this endocast is typical of ape brains but rarely, if ever, seen in human brains¹⁴: the configuration of TS and P, the configuration of IP and L, location of L well rostral to the lambdoid suture and a cerebellum that projects further posteriorly than the occipital pole. The latter two features indicate that there is no evidence either for expansion of parietal association cortices or for reorganization of visual cortices beyond the ape-like condition in the Hadar fossil hominids. The cranial capacity of AL 162-28 was less than 400 cm³, which suggests that brain volume in Hadar hominids may have been significantly smaller than brain volume in later living australopithecines.

The oldest human-like sulcal pattern in the fossil record to date is reproduced on an endocast from the 2-Myr-old KNM-ER 1470 (*Homo habilis*)¹⁰ whose cranial capacity¹⁸ is over 750 cm³.

On the other hand, all of the available evidence indicates that although brain size increased in australopithecines, the external morphology of their brains remained ape-like in appearance¹⁹⁻²¹. This in turn suggests that earliest hominid brain evolution may have occurred more at subcortical than at cortical levels.

I thank D. Johanson for permission to study AL 162-28, W. Kimbel for copies of the endocast and calvaria, the Government of Kenya for permission to conduct research, R. Leakey for access to Kenya National Museum hominids, and A. Turner for discussion and criticism of the manuscript. Figure 1 was prepared by A. Montalvo. This research was supported by NSF grants BNS-8203764 and BNS-8318253.

Received 13 March; accepted 1 August 1984.

- Holloway, R. L. *Nature* **303**, 420-422 (1983).
- Holloway, R. L. *Nature* **227**, 199-200 (1970).
- Holloway, R. L. *Hum. Neurobiol.* **2**, 105-114 (1983).
- Kimbel, W., Johanson, D. & Coppens, Y. *Am. J. phys. Anthropol.* **57**, 453-499 (1982).
- Brothwell, D. R. *Digging Up Bones* (British Museum, London, 1972).
- Passingham, R. *New Scientist* **68**, 510-511 (1975).
- Falk, D. & Kasinga, S. *J. hum. Evol.* **12**, 515-518 (1983).
- Leakey, M. G. & Leakey, R. E. F. *The Fossil Hominids and an Introduction to Their Context, 1968-1974* (Clarendon, Oxford, 1978).
- Day, M., Leakey, R., Walker, A. & Wood, B. *Am. J. phys. Anthropol.* **45**, 369-436 (1976).
- Falk, D. *Science* **221**, 1072-1074 (1983).
- Radinsky, L. B. in *The Primate Brain, Advances in Primatology* (eds Noback, C. R. & Montagna, W.) 209-224 (Appleton-Century-Crofts, New York, 1970).
- Radinsky, L. B. in *The Functional and Evolutionary Biology of Primates* (ed. Tuttle, R.) 175-184 (Aldine, Chicago, 1972).
- Jerison, H. J. in *Primate Brain Evolution: Methods and Concepts* (eds Armstrong, E. & Falk, D.) 77-84 (Plenum, New York, 1982).
- Connolly, C. J. *External Morphology of the Primate Brain* (Thomas, Illinois, 1950).
- Tobias, P. V. *Olduvai Gorge Vol. 2* (Cambridge University Press, 1967).
- Tobias, P. V. *Sond. Anthropol. Humangenet.* **1**, 1-10 (1968).
- Falk, D. & Conroy, G. *Nature* **306**, 779-781 (1983).
- Holloway, R. L. in *Early Hominids in Africa* (ed. Jolly, C.) 379-401 (St Martins Press, New York, 1978).
- Radinsky, L. B. *The Fossil Record of Primate Brain Evolution* (James Arthur Lecture, American Museum of Natural History, New York, 1979).
- Falk, D. *Am. J. phys. Anthropol.* **53**, 525-539 (1980).
- Falk, D. *Am. J. phys. Anthropol.* **60**, 479-489 (1983).

Disruptive selection for alternative life histories in salmon

Mart R. Gross

Department of Biological Sciences, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6

Many salmon species include males which mature as much as 50% younger and as small as 30% of the adult body size of other males in the population¹⁻⁸. In the semelparous Pacific salmon *Oncorhynchus* spp., these small males are known as 'jacks', and they compete with larger late-maturing 'hooknose' males for opportunity to spawn on the breeding grounds. The existence of jacks is problematical as it is believed that salmon populations should have a single optimal age or size at maturity⁹⁻¹¹. I now show, however, that these two alternative life histories are evolutionarily favoured by frequency-dependent disruptive¹² selection on the breeding grounds. In coho salmon (*Oncorhynchus kisutch*), small and large males gain access to females by sneaking and fighting respectively. By contrast, intermediate-sized males are at a competitive disadvantage. Jacks, which are specialized at sneaking, and hooknose males, which are specialized at fighting, have negatively frequency-dependent fitnesses from these alternative breeding tactics. Calculations suggest that the lifetime fitness of jacks is similar to that of hooknose males. Thus, age of maturity in salmon has probably not evolved as a single optimum, but rather as a 'mixed evolutionarily stable strategy'^{13,14} in which precocious maturity is an evolutionarily viable alternative life history strategy.

In 1981-82 I studied a wild population of coho breeding in Deer Creek Junior, a small stream in Washington State, United States. This species was selected because males mature at only two ages (Fig. 1). Field procedures involved collecting fish at a weir near the mouth of the stream as they migrated upstream to spawn. Before release above the weir to spawn naturally,

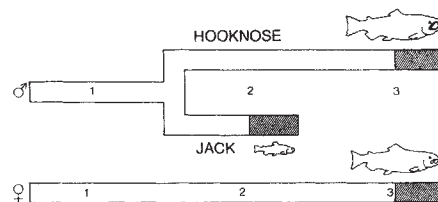


Fig. 1 Two life histories exist in male coho salmon: precocious maturity at age 2 as a small jack, or delayed maturity to age 3 as a large hooknose²⁶. The term 'hooknose' is derived from the exaggerated snout and enlarged teeth which develop at maturity and are used in fighting. By contrast, 'jacks' lack secondary sexual characters and are relatively cryptic on the breeding grounds. Both jack and hooknose males die after breeding, as do the females. Coho life history involves autumn-winter spawning (November-January); egg and larval incubation (November-March) in gravel nests excavated by females; 1 yr stream residency as fry before migration to the ocean as smolts; 5-8 months' ocean residency before return as jacks, or 17-20 months before return as mature females and hooknose males. Hatched areas indicate reproductive maturity.

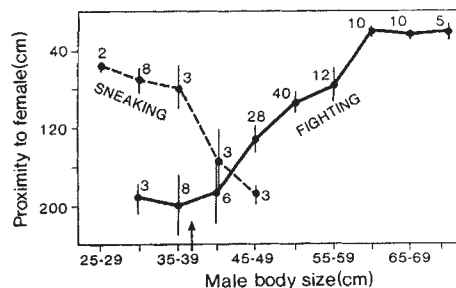


Fig. 2 This plot shows a threshold in male size above which female proximity is best gained by fighting, and below by sneaking. The success of male coho on the breeding ground depends on their body size and behavioural tactic. Proximity to ovipositing females may be obtained through two tactics: sneaking or fighting.

Methods: Behavioural and proximity data are from naturally breeding males of known body size (sample sizes and standard errors are shown). Two techniques were used to measure distance from the nest centre to the mid-region of the male: (1) male position was marked by eye, and immediately after spawning the distance was measured using a tape measure, or (2) the distance was estimated by eye, with estimates having an error of $\pm 6\%$ based on 19 repeated measurements. These two techniques were used randomly for fishes of all sizes and behavioural tactics. The body size range of hooknose males (fork length) was 39-72 cm (mean = 51.7 ± 4.2), that of jack males 25-39 cm (mean = 33.6 ± 4.7). The size distinction between jack and hooknose males is indicated on the size axis with an arrow. The proximity of hooknose males to females achieved through fighting averaged 90 cm ($n=111$ observations) and through sneaking 170 cm ($n=6$). By contrast, jack proximity achieved through fighting averaged 197 cm ($n=11$) and through sneaking 71 cm ($n=13$). These proximities are statistically different within male type, and between male types (Mann-Whitney U tests, $P < 0.05$ for the four comparisons).

their sex was determined, and body length measured, and each fish was tagged with colour-coded Petersen disks. Tagging permitted subsequent recognition of individuals by observation from the stream bank without disturbing spawning.

Thirty-seven randomly chosen breeding groups were studied, each of which included a sexually active female and several males. While the female excavated a nest for oviposition, the males competed among themselves for proximity to the nest. Proximity is important in determining male mating success, because successive males fertilize smaller proportions of eggs¹⁵ and nest entry order is strongly correlated with male proximity ($r = 0.73$, $n = 66$ males, $P < 0.001$). Males make use of two tactics to gain proximity: (1) fighting, where a linear dominance hierarchy is established with the 'alpha' male closest to the female and subordinates more distant; and (2) sneaking, where males use 'refuges' (rocks, debris, or shallow areas in the stream)

to escape aggression from larger males while remaining close to the nest site. Male success from fighting and sneaking was size-dependent: larger males were best at fighting; smaller males were best at sneaking. Males of intermediate size could not obtain suitable refuges for sneaking, nor could they fight successfully against the larger males. These males were usually most distant from the female (Fig. 2). During mating, natural selection thus favours large- and small-bodied males at the expense of intermediate-sized males, thereby placing body size under disruptive selection.

The small, early maturing jack males were ideally suited for sneaking (Fig. 1), but poorly suited for fighting. Opportunity to sneak depended on the availability of suitable refuges, which were limited in quantity and could generally be used by only one jack at a time. Since jack density in the stream was too high for all to sneak successfully, some were forced to fight. In 45% of jack matings, jacks had to fight for proximity and obtained positions at the low end of dominance hierarchies (Fig. 2). The average proximity of jacks from the 'mix' of sneaking and fighting observed was 124.6 ± 15.5 cm (mean $\pm$ s.e., $n = 24$). By contrast, the average proximity of hooknose males was 93.0 ± 6.1 cm ($n = 117$; Mann-Whitney, $z = 1.96$, two-tailed $P = 0.05$). Because the relationship between proximity and success is approximately 'fertilization success = a constant/male proximity' (ref. 15), relative male proximity provides a means of measuring relative male mating success. Thus, assuming random sperm competition among males, a reasonable assumption because of the nature of this mating system¹⁶, jack success was approximately 66% ($1 - [124.6 - 93.0]/93.0$) that of hooknose males. Any increase in the proportion of jacks would decrease their average proximity because more jacks would be forced to fight; conversely, any decrease in jack frequency would increase average proximity because relatively more jacks could sneak matings. Thus, the mating success of jacks relative to hooknose males is negatively frequency-dependent.

Negatively frequency-dependent disruptive selection may, in theory, give rise to stable alternative life history strategies^{13,17,18}. If the alternative life histories are a mixed evolutionarily stable strategy (mESS) in salmon, the lifetime fitnesses of jack and hooknose males should be equal at the frequency observed. In the coho population, fitness calculations are relatively simple since the population has discrete generations and, to the best of my knowledge, is stationary¹⁹. Lifetime fitness can be estimated from the product of probability of survival to maturity, breeding lifespan, and mating success. In calculating survival to maturity, I assume that early freshwater survival of young jacks and hooknose males is equal, and that the critical difference in survivorship is ocean mortality. The study population is part of the Skykomish River coho stock for which the Washington State Department of Fisheries maintains ocean mortality records²⁰. In this stock the average ocean survivorship is 13% for jacks and 6% for hooknose males. I have found from following tagged individuals on the breeding grounds that jacks are reproductively active for 8.4 ± 2.3 days ($n = 7$), and the hooknose males for 12.7 ± 1.2 days ($n = 35$). I assume that the number of opportunities for spawning is proportional to length of life on the breeding grounds. Therefore, the relative lifetime fitness of jack (W_j) and hooknose (W_h) life histories is:

$$\begin{aligned} \frac{W_j}{W_h} &= \frac{0.13}{0.06} (\text{survivorship to maturity}) \\ &\times \frac{8.4}{12.7} (\text{breeding lifespan}) \\ &\times \frac{0.66}{1} (\text{mating success}) \\ &= 0.95 \end{aligned}$$

This calculation suggests that the fitnesses of the alternative life histories approach equality. The estimate may be conservative for jacks as it does not include fishing pressure, which is greater on hooknose males than jacks. How salmon populations of the

west coast are responding evolutionarily to fishing pressure is not yet known²¹.

The similarity in lifetime fitness estimates combined with evidence for negatively frequency-dependent success suggests that the jack life history exists in a mixed evolutionarily stable strategy with hooknose males. This conclusion is supported by two findings: (1) Those coho males maturing as jacks are not smaller in body size at age 2 (refs 5, 22) and thus are not competitively inferior to males that delay maturity to become hooknoses. A size difference would be expected if jack maturity was a BBS strategy (best of a bad situation)²³ rather than a mESS. (2) Recent breeding studies suggest that jack maturity is heritable²⁴. Heritability in a stationary population will probably exist only through equal fitnesses, as genes coding for an inferior life history should be removed from the gene pool by natural selection^{18,25}.

This work was supported by a NSERC of Canada Operating Grant and a NSERC University Research Fellowship. I thank S. Rohwer, E. van den Berghe, B. Jonsson, E. Charnov and K. Reidinger for discussion, N. Gerrish, E. van den Berghe, R. Peterman, L. Dill and C. Sargent for comments on the manuscript, and E. van den Berghe for valuable field assistance.

Received 20 July; accepted 9 October 1984.

1. Jones, J. W. *The Salmon* (Harper, New York, 1959).
2. Geghards, S. V. *Prog. Fish. Cult.* **22**, 121-123 (1960).
3. Kimura, S. *Jap. J. Ichthyol.* **19**, 111-119 (1972).
4. Ricker, W. E. in *The Stock Concept in Pacific Salmon* (eds Simon, R. C. & Larkin, P. A.) 19-160 (University of British Columbia, Vancouver, 1972).
5. Hager, R. C. & Noble, R. E. *Prog. Fish. Cult.* **38**, 144-147 (1976).
6. Schmidt, S. P. & House, E. W. *J. Fish. Res. Bd Can.* **36**, 90-93 (1979).
7. Jonsson, B. thesis, Univ. Oslo (1981).
8. Maekawa, K. *Jap. J. Ichthyol.* **30**, 227-234 (1983).
9. Schaffer, W. M. & Elson, P. F. *Ecology* **56**, 577-590 (1975).
10. Schaffer, W. M. *Symp. zool. Soc. Lond.* **44**, 307-326 (1979).
11. Crandall, R. E. & Stearns, S. C. *Theor. Populat. Biol.* **21**, 11-23 (1982).
12. Thoday, J. M. *Proc. R. Soc. B182*, 109-143 (1972).
13. Maynard Smith, J. *Evolution and the Theory of Games* (Cambridge University Press, 1982).
14. Gross, M. R. *Z. Tierpsychol.* **60**, 1-26 (1982).
15. Schroder, S. L. in *Salmon and Trout Migratory Behavior Symp.* (eds Brannon, E. L. & Salo, E. O.) 275-285 (School of Fisheries, University of Washington, Seattle, 1982).
16. Kazakov, R. V. *J. Fish. Biol.* **18**, 1-8 (1981).
17. Gross, M. R. & Charnov, E. L. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6937-6948 (1980).
18. Charnov, E. L. *The Theory of Sex Allocation* (Princeton University Press, 1982).
19. Charlesworth, B. *Evolution in Age-Structured Populations* (Cambridge University Press, 1980).
20. 1979-1983 Stock Input Information (Washington State Department of Fisheries, Washington DC, 1984).
21. Ricker, W. E. *Can. J. Fish. aquat. Sci.* **38**, 1636-1656 (1981).
22. Bilton, H. T., Alderdice, D. F. & Schnute, J. T. *Can. J. Fish. aquat. Sci.* **39**, 426-447 (1982).
23. Gross, M. R. in *Fish Reproduction: Strategies and Tactics* (eds Potts, G. W. & Wootton, R. J.) 55-75 (Academic, London, 1984).
24. Iwamoto, R. N., Alexander, B. A. & Hershberger, W. K. *Salmonid Reproduction*, 8 (School of Fisheries, University of Washington, Seattle, 1983).
25. Slatkin, M. *Am. Nat.* **114**, 384-398 (1979).
26. Marr, J. C. *Stanford Ichthyol. Bull.* **2**, 157-197 (1943).

Visual control of the partition of flight force between lift and thrust in free-flying *Drosophila*

C. T. David

AFRC Insect Physiology Group, Department of Pure and Applied Biology, Imperial College of Science and Technology, London SW7 2AZ, UK

Flying fruit flies have a system for stabilizing height and speed which has two main components. They can change the total force generated by their wings and they can adjust the division of this force between lift and thrust by changing the angle of their body axis to the horizontal¹⁻³. The optomotor reactions determining the force from the wings are governed by the vertical components of the image flow velocity^{2,4-6}. I show here that the partition between lift and thrust is determined by the horizontal components of the image flow velocity. The resolution of the image flow velocity into components at right angles, each controlling a different output, produces a simple height and speed stabilizing system.

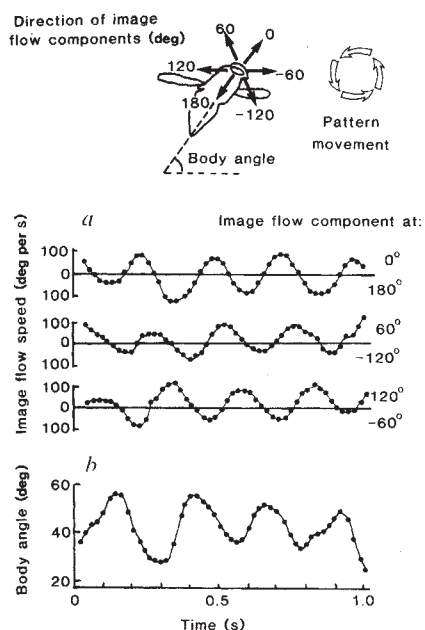


Fig. 1 *a*, The image flow speed components at various directions relative to the body axis of a fruit fly; *b*, the body angle of the fly. The pattern was moved at 4 Hz in a circular path in an anticlockwise direction without rotation. The image velocity in each direction is approximately sinusoidal, as shown by the three pairs of coupled speed curves. The speed curve in each direction is shaped like a series of half-sinusoids, each delayed slightly compared with those of the previous direction, and separated by regions of 'negative' speed.

Drosophila hydei Sturtevant were flown in a horizontal transparent tube 15 cm in diameter against a smooth air flow of 40 cm s^{-1} . On either side of and parallel with this tube were transparent vertical screens 40 cm high by 80 cm long and 30 cm apart, bearing a number of irregularly shaped, transparent orange spots 2–5 cm across and about 2 cm apart. These screens were moved bodily through a circular path without rotation (like a bicyclist's foot) in a vertical plane parallel to the transparent tube, using cranks with a throw of 2.54 cm at 2–4 revolutions per s. By this means, flies flying inside and along the transparent tube were presented on each side with pattern movements sequentially in each direction. The component of the image velocity in each direction varied approximately sinusoidally during this process, but out of phase with the velocity components in any other direction. Insect movement detectors respond to the component of the visual flow velocity in one direction—that is, they respond to the speed of flow in one direction along an axis and not to the speed of flow in the other direction—and the detectors for different directions of movement are not necessarily opposite to one another⁵. Because of this, the image speeds in each direction were used in the analysis rather than the image velocities along particular axes (Fig. 1).

The speeds were calculated as the relative angular velocity of the point on the screens closest to the fly. The entire apparatus was surrounded by a white screen and illuminated from outside by d.c. fluorescent lamps. The flies were video-recorded from the side as they flew upwind, using a camera equipped with a rotating shutter giving an exposure of $1/1,000 \text{ s}$ every $1/50 \text{ s}$. Only flies that stayed in the field of view of the camera for at least 1 s and in the midline of the tunnel without turning were included in the analysis. The position of a fly, a point on an extension of the line passing through its long body axis, and a point on the moving pattern were digitized from each frame of the video record. From these data, the fly's body angle and 36 of the speed components of the relative movement between it and the pattern (the image flow speed over lateral-facing ommatidia for every 10° relative to its body axis) were calculated.

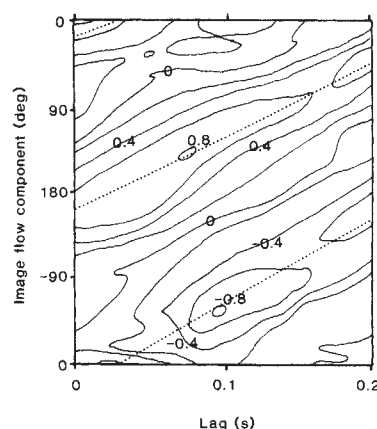


Fig. 2 Contour map of the lag correlations between the image flow speed resolved into 36 directions and the subsequent body angle of a fly, where the pattern was moved anticlockwise (as in Fig. 1), but at 2 Hz. The contour interval is 0.2. The oblique dotted lines mark the lines of positive and negative correlation peaks that were superimposed with others to form Fig. 3.

Figure 1 shows the body angle of one fly and six of the components of the image flow speed for 1 s. The body angle changed in time with the image flow components and must therefore have been determined by at least one of them. The flies' reaction time was unknown, so it is difficult to determine which of the components of the image flow speed drove the body angle because they all followed the same approximately half-sinusoidal time course. Thus, the body angle might have been controlled either by the component at 0° to the fly's body axis with a lag (reaction time) of 0.2 s, or by the component at 120° to the body axis with a lag of 0.09 s.

Because the image flow field rotates almost uniformly relative to a fly's body, a plot of the directions of the possible controlling image speed components against their corresponding lag times will be nearly linear. This linear relationship will be reversed if the image flow is rotated in the opposite direction. The intersection obtained by superimposing these two results gives the only values of the direction of the speed component and lag time which are consistent with both motions. The plot of the directions of the image speed component and the corresponding lag times were obtained in the following way. Contour maps were drawn of the correlation coefficients between the body angle and each of the 36 components of the image flow speed at lags of 0–0.2 s. The portions of negative image flow speed (where the image flow had no component in that direction) were omitted from the correlations. There is a line of strongly positive correlations with decreasing lag as the image flow component changes from 0° to 180° and similarly of strongly negative correlations from 180° on to 0° (Fig. 2). The correlation peaks on the contour map are the plot of the possible controlling image speed components and the corresponding lags described above (Fig. 2).

When the pattern was moved in the opposite direction, the maps, as expected, also showed lines of correlation, but with increasing lag as the image flow component changed from 0° to 180° and 180° on to 0° . In superimposed plots (Fig. 3), the points where the lines of like correlation peaks cross must mark the particular components and lags of the image flow speed which controlled the body angle. These values were, on average, at 130° and -50° and with a lag of 0.08 s (Fig. 3); the ranges were $\pm 20^\circ$ and 0.03 s. Also, it is the body angle that is controlled by these image flow components and not the change in this angle (see Fig. 3 legend).

The points where the lines of correlation peaks cross were thus 180° apart, showing that the components of image flow affecting the body angle were 180° apart. This differs from the components of image flow which affect the total force produced by the wings, for these are $\sim 120^\circ$ apart⁵.

These experiments show that *Drosophila* have a response that controls the way in which the force generated by their wings is

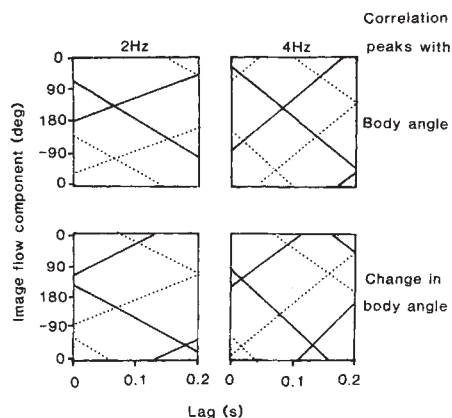


Fig. 3 Plot of the directions of the possible controlling image speed components and their corresponding lags, derived by superimposing the average position of the lines of peaks of the lag correlations as in Fig. 2, for 20 flies, 10 with the pattern moved clockwise and 10 with the pattern moved anticlockwise at 2 Hz, and for 20 flies treated similarly, with the pattern moved at 4 Hz. Continuous lines, positive peaks; dashed lines, negative peaks. The crossing points of the lines of like correlation peaks show the angle relative to the body axis and the lag where a component of the image flow speed controls the body angle or the change in the body angle. The lags of the crossing points differ when the pattern is moved at 2 Hz and at 4 Hz in the plot for the change in the body angle, but do not do so in the plot of body angle. This indicates that it is the body angle which is controlled by the image flow speed at 130° and -50° to the body axis and not the change in the body angle, because the reaction time (and hence the lags) is unlikely to be affected by this change in the frequency of pattern movement⁶. The changes in lag in the plots for change in body angle are expected as a consequence of the differentiation of the various frequencies of the sinusoidally changing body angle, and are not a physiological process.

balanced between lift and thrust in response to movement past their eyes at 130° or -50° . This is nearly horizontal movement, as the fly's average body angle in these experiments at a wind speed of 40 cm s^{-1} was 43° . If a fly's groundspeed becomes too high, the backwards image flow speed becomes higher than some 'set point'; the body angle is then increased, reducing the fly's thrust and hence decelerating it; and vice versa. *Drosophila* also have a response which controls the total force from the wings. The image flow components which control this vary so as to remain vertical relative to the outside world as the body angle changes with wind speed⁷. Therefore, the components of the image flow which control the total force from the wings are at right angles to the components which control the partition between lift and thrust. A similar orthogonal arrangement of sensitive directions to image flow speed has been found in tethered flies; this is between the optomotor turning reaction and that controlling the total force from the wings². It has also been found electrophysiologically⁸⁻¹⁰. Together, these visual feedbacks provide an elegant system that compensates almost completely for any errors that may occur in a fly's speed and height as a result of changes in wind speed or air turbulence.

I thank J. S. Kennedy, C. Rechten and V. A. Taylor for their helpful comments on the manuscript.

Received 29 August; accepted 16 October 1984.

1. Vogel, S. J. *exp. Biol.* **44**, 567-578 (1966).
2. Götz, K. G. *Kybernetik* **4**, 199-208 (1968).
3. David, C. T. *Physiol. Ent.* **3**, 191-195 (1978).
4. Buchner, E., Götz, K. G. & Straub, C. *Biol. Cybernet.* **31**, 235-242 (1978).
5. Götz, K. G. & Buchner, E. *Biol. Cybernet.* **31**, 243-248 (1978).
6. David, C. T. *Physiol. Ent.* (in the press).
7. Götz, K. G. *Drosoph. Inf. Serv.* **59**, 39-40 (1983).
8. McCann, G. D. & Foster, S. F. *Kybernetik* **8**, 193-203 (1971).
9. Hausen, K. *Biol. Cybernet.* **45**, 143-156 (1982).
10. Eckert, H. J. *comp. Physiol.* **149**, 195-205 (1982).

Rhodopsin-like sensitivity of extra-retinal photoreceptors mediating the photoperiodic response in quail

R. G. Foster, B. K. Follett & J. N. Lythgoe

AFRC Research Group on Photoperiodism & Reproduction, Department of Zoology, University of Bristol, Bristol BS8 1UG, UK

It has been known for some 50 years that birds use photoreceptors in or near the hypothalamus to mediate the photoperiodic responses that control seasonal breeding¹. So far, however, attempts to identify the photopigment by determining an action spectrum have failed¹. The problems stem from the selective filtering of light by the tissues surrounding the photoreceptors and the need to deliver defined amounts of light over the days or weeks required to induce a quantitative measure of photostimulation. Here we have developed a technique which produces a quantitative action spectrum for the photoperiodic response in the Japanese quail; the results indicate the presence of a rhodopsin photopigment with a peak sensitivity of $\sim 492 \text{ nm}$. The photoreceptors exhibit a level of sensitivity comparable with that of vertebrate visual pigments. We conclude that the brain photoreceptors of birds are based on a rhodopsin/rhodopsin-like photopigment.

Removing the eyes and/or the pineal gland from Japanese quail does not influence the effects of 'long days' on gonadal maturation^{2,3}. Specifically, illuminating the pineal gland has no effect on the gonads⁴ but local illumination of the posterior basal hypothalamus leads to gonadal growth, as if the birds had been exposed to long days¹. Lesions in this hypothalamic region block totally the photoperiodic response⁵. Previous studies on the spectral sensitivity of the brain photoreceptors suggested a red-sensitive photopigment¹ but as red light passes through the brain much more effectively than shorter wavelengths⁶, these studies may have given a false impression as to the true nature of the pigment.

In the present study, monochromatic light (15 nm half-peak bandwidth) was obtained by passing light from a voltage-stabilized 100 W quartz-halogen lamp through a monochromator (Oriel Corporation). A lens assembly split the beam into three glass optic fibres (diameter 3 mm, length 400 mm); the light intensity leaving the end of each fibre was measured with a spectroradiometer (Macam Photometrics) calibrated in $\mu\text{E cm}^{-2} \text{ s}^{-1} \text{ nm}^{-1}$ ($1 \mu\text{E} = 6 \times 10^{17}$ photons). This light was used to photostimulate Japanese quail. One fibre optic was attached to the surface of the bird's skull and secured by a permanently mounted head socket. The triple-fibre optic allowed three birds to be photostimulated simultaneously. During illumination, each bird was kept in a cage screened from extraneous light and sound.

To obtain the action spectrum, we first measured the absorption of different wavelengths of light as they passed through the skull and brain tissues⁷ of recently dead and deeply anaesthetized birds, both of which gave identical results. White light from a quartz-halogen lamp was delivered to the skull via a fibre optic and its transmission measured at the exposed surface of the basal hypothalamus (12 birds) or the parolfactory area (10 birds). As the light from the quartz-halogen lamp does not contain the same number of photons at all wavelengths, transmission was expressed as the fraction of transmission which would occur if no tissue was placed between the light and the detector. The spectral composition of the light reaching the hypothalamus and parolfactory region was not significantly different, the transmission pattern being dominated by the absorption characteristics of haemoglobin. Thus, transmission of wavelengths of $\sim 500 \text{ nm}$ was 30 times less than that of wavelengths of $\sim 690 \text{ nm}$. The transmission curves were used to devise a multiplication factor at each wavelength that would compensate for the relative absorption of that wavelength as it passed through the brain. These factors were then used to adjust

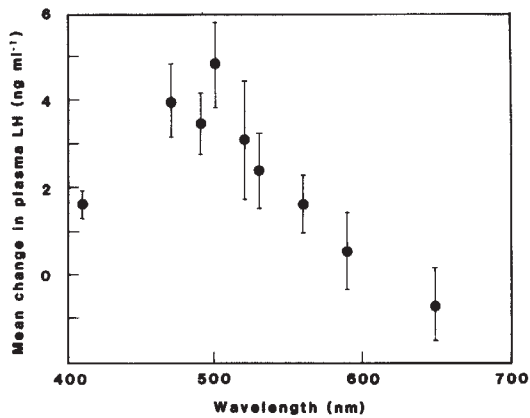


Fig. 1 The effect of light of differing wavelength but equal numbers of photons on the photoperiodic response in Japanese quail. Each point shows the mean ($\pm$ s.e.m.) increase in the concentration of plasma LH after exposure of the brain photoreceptors to a single 'long day' of monochromatic light. At least nine birds were used for each point (range 9–16). The photoreceptors appear most sensitive to light at wavelengths of ~ 500 nm.

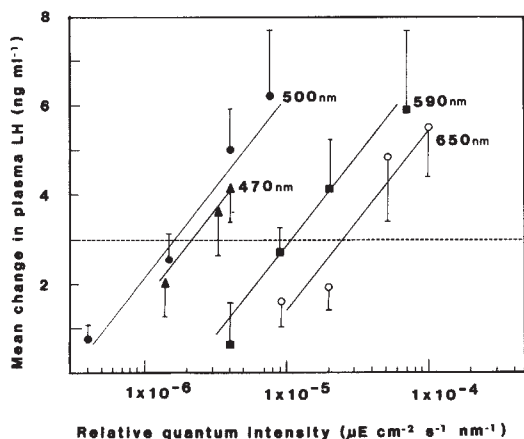


Fig. 2 Dependence of the photoperiodic response of quail on the intensity of light applied to the extraretinal photoreceptors. Results shown for four wavelengths: 470, 500, 590 and 650 nm. Each point represents the mean ($\pm$ s.e.m.) change in plasma concentration of LH. The number of quail used at each point ranged from 10 to 22. A regression line has been fitted to the data points at each wavelength. The photon flux (relative to that at the skull surface at 500 nm) is given on the abscissa. Adjustments were made for the other wavelengths to ensure that absorption through the skull and brain were taken into account. The broken line shows the light intensity used to obtain an equal photoperiodic response (see Fig. 3). Light at 500 nm is the most effective at inducing a response.

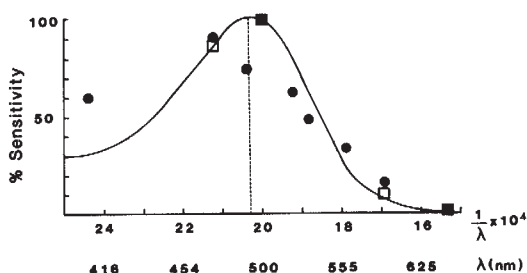


Fig. 3 Combined action spectrum for the photoperiodic response in Japanese quail, derived from Figs 1 (●) and 2 (□); ■ indicates where the data give identical results. Using Dartnall's correction¹¹, the data have been plotted on the basis of photons per equal-frequency interval ($1/\lambda$) so that they may be directly compared with a visual pigment 'nomogram' curve. A curve for rhodopsin fits the data points well, a rhodopsin with a maximum absorption at 492 nm providing the best fit.

the intensity of light delivered at the skull surface so that equal relative numbers of photons arrived at the hypothalamus. The spectral transmission of light into more dorsally located areas of the brain, such as the tectum, is also dominated by the absorption spectrum of haemoglobin. However, beyond wavelengths of 620 nm the hypothalamic area showed a relative enrichment in the amount of long-wavelength light reaching this region. At 690 nm the hypothalamus received, relatively, three times more light than the tectum; this discrepancy was not seen at shorter wavelengths. If a photoreceptor did lie very dorsally in the brain then our action spectrum would only become distorted beyond 620 nm. In fact, as mentioned above, there is no evidence for a tectal or pineal photoreceptor being involved in photoperiodism.

To determine whether any of the light delivered via the fibre optic to the skull surface reached the retina, bright white light ($10^{-2} \mu\text{E cm}^{-2} \text{s}^{-1}$, integrated between 450 and 750 nm) was used to illuminate the brain of dead birds ($n=6$) and the detector head of the spectroradiometer placed in the posterior chamber of the eye; even at this light intensity, about three orders of magnitude above that encountered on a bright sunny day⁸, no light could be detected in the eye. Clearly, the black pigmented layer behind the retina shields it from any back-illumination. Therefore, there was no danger of the illumination directed at the brain reaching the retina.

To measure photoperiodic induction we used castrated male quail, which, when exposed just once to a photoperiod of 20 h, produce a quantifiable and repeatable increase in secretion of luteinizing hormone (LH)⁹. In these experiments, we used a batch of 55 castrated male quail, each of which was fitted with a small head socket by which it could be connected to the fibre optic. Groups of three quail were photostimulated daily, with an individual bird being used at monthly intervals. Just before 'dusk' on the experimental day, quail were attached to the fibre optics and exposed to a further 12 h of light (making the equivalent of 20 h light/4 h dark). During this period, different intensities and wavelengths of light were applied to the brain photoreceptors. At 20 h the birds were returned to their normal cages. Blood samples (200 μl) were collected the day before and 3 days after photostimulation and concentrations of LH measured by radioimmunoassay¹⁰. The change in plasma LH concentration (ng ml^{-1}) was used as a measure of photoperiodic induction.

The first action spectrum used a single light intensity and determined the LH responses at nine wavelengths (Fig. 1). The actual intensity of light delivered to each bird's skull was $4 \times 10^{-6} \mu\text{E cm}^{-2} \text{s}^{-1}$ at 500 nm; at other wavelengths it was increased or decreased by an amount proportional to the absorption of that wavelength relative to 500 nm. A second series of experiments measured the responses at four wavelengths (470, 500, 590 and 650 nm) as a function of light intensity. The parallel nature of these curves (Fig. 2) suggests that only a single photo-sensitive step is involved. By taking the reciprocal of the light intensity required to induce an equal photoperiodic response (derived from the broken line in Fig. 2) and by normalizing the data with those in Fig. 1, the two action spectra can be combined (Fig. 3). Using Dartnall's method for fitting visual pigment curves¹¹, we obtained a best fit between his standard curve and our experimental action spectrum (see Fig. 3); a rhodopsin with a maximum absorption at 492 nm provided the best fit.

By determining how much light is actually absorbed across the quail's skull and brain, and relating this to the levels of light necessary to trigger a significant photoperiodic response ($4 \times 10^{-7} \mu\text{E cm}^{-2} \text{s}^{-1}$ at 500 nm), an estimate can be made of the absolute sensitivity of the brain photoreceptors. When light at 500 nm is applied to the skull surface, the intensity is reduced by a factor of 1.4×10^5 before it reaches the basal hypothalamus so that the absolute sensitivity at this wavelength must be $4 \times 10^{-7} / 1.4 \times 10^5 = 2.85 \times 10^{-12} \mu\text{E cm}^{-2} \text{s}^{-1}$, or $\sim 2 \times 10^{10}$ photons $\text{m}^{-2} \text{s}^{-1}$, a level of sensitivity matched only by the photopigments of higher vertebrates¹².

In the wild the feathers on the head would play only a minor part in modifying the spectral composition of light reaching the brain; they act rather like a neutral density filter, reducing the total light intensity by a factor of ~ 60 .

Although the brain photoreceptors appear to be based on a rhodopsin photopigment most sensitive to blue-green light, their environment, because of haemoglobin absorption, is dominated by red light. This mismatch is not restricted to the brain photoreceptors as many rod rhodopsins in the vertebrate visual system also absorb most strongly at shorter wavelengths than one would expect from the spectral composition of their light environment¹³. Why this should be the case is uncertain, but as photoreceptive cells act as photon counters they can, within limits, compensate for a low photon flux by collecting photons over a longer period¹⁴. It is possible that the brain photoreceptors of birds achieve their surprising sensitivity by using a relatively long integration time—this hypothesis is supported to some extent by evidence in the literature^{15,16}.

Received 9 May; accepted 22 October 1984.

1. Oliver, J. & Bayl , J. D. *Experientia* **38**, 1021–1029 (1982).
2. Siopes, T. D. & Wilson, W. O. *Poultry Sci.* **53**, 2035–2041 (1974).
3. Simpson, S. M. & Follett, B. K. *J. comp. Physiol.* **144**, 381–389 (1981).
4. Homma, K., Ohta, M. & Sakakibara, Y. in *Biological Rhythms in Birds* (eds Tanaka, Y., Tanaka, K. & Ookawa, T.) 149–156 (Springer, Berlin, 1980).
5. Davies, D. T. & Follett, B. K. *Proc. R. Soc. B* **191**, 285–301 (1975).
6. Hartwig, H.-G. & van Veen, T. J. *comp. Physiol.* **130**, 277–282 (1979).
7. Knowles, A. & Dartnall, H. J. A. *The Photobiology of Vision* (Academic, New York, 1977).
8. McFarland, W. N. & Munz, F. W. in *Light as an Ecological Factor* Vol. 2 (eds Evans, G. C., Bainbridge, R. & Rackham, O.) (Blackwell Scientific, Oxford, 1975).
9. Nicholls, T. J., Follett, B. K. & Robinson, J. E. *J. Endocr.* **97**, 121–126 (1983).
10. Follett, B. K., Scanes, C. G. & Cunningham, F. J. *J. Endocr.* **52**, 359–378 (1972).
11. Dartnall, H. J. A. in *Vision in Fishes* (ed. Alis, M. A.) 543–563 (Plenum, New York, 1975).
12. Land, M. F. in *Handbook of Sensory Physiology* Vol. 7/6B, Ch. 4, 472–532 (1980).
13. Lythgoe, J. N. in *Environmental Physiology of Fishes* (ed. Ali, M. A.) (Plenum, New York, 1980).
14. Lythgoe, J. N. *The Ecology of Vision* (Clarendon, Oxford, 1979).
15. Follett, B. K. & Millette, J. J. *J. Endocr.* **93**, 83–90 (1982).
16. Ellis, D. H. & Follett, B. K. *Biology of Reproduction* **29**, 805–818 (1983).

The apical surface of canine chief cell monolayers resists H^+ back-diffusion

Martin J. Sanders*, Amram Ayalon†, Michal Roll† & Andrew H. Soll*

* Department of Medicine and the Center for Ulcer Research and Education, UCLA School of Medicine and the Medical and Research Services, VA Wadsworth Medical Center, Los Angeles, California 90073, USA

† Department of Surgery B, Hadassah University Hospital, Jerusalem, Israel

The resistance of the gastric mucosa to acid and peptic injury is reflected by a resistance to the back-diffusion of H^+ from gastric lumen to blood^{1,2}. The nature of this 'barrier', however, remains undefined. Using Ussing chambers, we have now studied the acid-barrier function of monolayers prepared from dispersed canine fundic chief cells³. These monolayers secrete pepsinogen in response to stimulation⁴. We found that, on acidification of the apical solution to pH 2, transepithelial resistance (R) increased 2.6-fold and the monolayers maintained this 1:100,000 H^+ concentration gradient for more than 4 h. The addition of aspirin to the acidified apical solution caused a rapid decay in R , as did acidification of the basolateral solution to a pH <5.5. Ouabain-treated monolayers displayed the rise in R expected with apical acidification, while potential difference (V) and short-circuit current (I_{sc}) decreased essentially to zero, indicating impermeability to H^+ . However, if the integrity of the ouabain-treated monolayers was disrupted by low apical pH, H^+ permeation occurred, reflected by an I_{sc} that was dependent on the H^+ gradient across monolayers. These data indicate that the apical surface of chief cells is a very tight barrier to H^+ diffusion and may be an important element resisting acid-peptic injury.

Canine chief cell monolayers, mounted in Ussing chambers, maintained electrical integrity (stable R) despite very large H^+

gradients (Fig. 1a). Once stabilized with an apical pH of 2–2.5, the pH in the unbuffered basolateral solution remained unchanged for at least 4 h. There was a critical apical pH below which the barrier to acid was overcome, causing a rapid decay in R . This critical pH value was between 2 and 2.5 in 70% of 60 monolayers tested and <2 in 20%. Exposure of monolayers to pepsin ($800 \mu\text{g ml}^{-1}$) for 2 h at an apical pH of 2.5 had no effect on R or I_{sc} ($n=3$). However, the addition of 4 mM aspirin (dissolved in dimethyl sulphoxide) to an acidified apical solution produced an immediate fall in R (Fig. 1a). Vehicle alone (1% dimethyl sulphoxide) or aspirin at neutral pH had no effect on either R or I_{sc} .

Acidification of the apical solution caused marked changes in R , I_{sc} and V , which occurred in two phases. In the range of apical pH between 6.5 and 3.5, R increased from $1,650 \pm 139$ to $2,450 \pm 170 \Omega \text{ cm}^2$ (mean $\pm$ s.e., $n=24$) without a significant change in I_{sc} (17 ± 2 to $16.2 \mu\text{A cm}^{-2}$; Fig. 1b, c). V consequently increased from 25 ± 2 to 34 ± 3 mV. At apical pH values below 3.5, R increased further to $3,600 \pm 178 \Omega \text{ cm}^2$ and I_{sc} decreased to $6 \pm 1 \mu\text{A cm}^{-2}$. In the latter phase, these opposing changes in R and I_{sc} resulted in a decrease in V to 19 ± 2 mV. The changes in R , I_{sc} and V were reversible on apical neutralization and could be reproduced repeatedly in the same monolayer on re-acidification of the apical solution (Fig. 1b). The return of transport function (I_{sc}) to control levels on apical neutralization indicated preservation of cellular integrity during acidification.

In contrast, acidification of the basolateral solution to below pH 5.5 produced a transient increase followed by a rapid decay in R (Fig. 1c). V fell with basolateral acidification, reflecting a profound decrease in I_{sc} .

The barrier function of epithelia represents the permeability to a solute via paracellular junctions and/or transcellular channels^{5–7}. Our finding that apical acidification produced two phase changes in R , I_{sc} and V suggests interaction of H^+ with both of these pathways. In the first phase (pH range 6.5–3.5), apical acidification produced no change in I_{sc} and, as I_{sc} reflects transcellular ion transport, the rise in R in this pH range probably resulted from proton effects on the paracellular shunt pathway. In the second phase (pH <3.5), I_{sc} decreased markedly, indicating that proton interaction with transcellular channels was responsible for the increase in R . We have shown previously that amiloride-sensitive Na^+ transport accounts for 70% of I_{sc} in chief cell monolayers³ and it is well known that Na^+ entry through amiloride-sensitive, transcellular channels is pH dependent^{5,8}.

Basolateral application of ouabain (1 mM) caused a marked decrease in I_{sc} (from 18 ± 5 to $2 \pm 1 \mu\text{A cm}^{-2}$) and V (from 24 ± 5 to 4 ± 1 mV) (Fig. 1d), but subsequent acidification of the apical solution still produced biphasic changes in R ; R increased 2.3-fold (from $1,720 \pm 450$ to $3,890 \pm 670 \Omega \text{ cm}^2$, $n=4$) and reached the same level achieved on apical acidification without ouabain. At pH <3.5, the small residual I_{sc} and V were abolished. These changes in R , I_{sc} and V were also reversible in ouabain-treated monolayers on apical neutralization and could be reproduced in the same monolayer on re-acidification of the apical solution. Monolayer integrity, as reflected by R , was maintained at an apical pH of 2.5. Thus, ouabain did not interfere with the ability of these monolayers to maintain a pH gradient nor prevent the rise in R with apical acidification, indicating that this barrier to acid back-diffusion is independent of electrogenic ion transport. Furthermore, these chief cell monolayers displayed barrier function in the absence of basolateral bicarbonate. Thus, our data provide no support for the hypothesis that barrier function depends on either Na^+ or Cl^- transport^{5,9}, but instead seems to reflect a passive property of the apical membranes and tight junctions of these gastric mucosal cells.

With apical acidification, the potential difference of ouabain-treated monolayers was abolished, despite the 1:100,000 concentration gradient of H^+ ions (apical pH 2.5 versus basolateral pH 7.4). Such a concentration gradient would generate a potential difference in a system permeable to H^+ , and the absence of

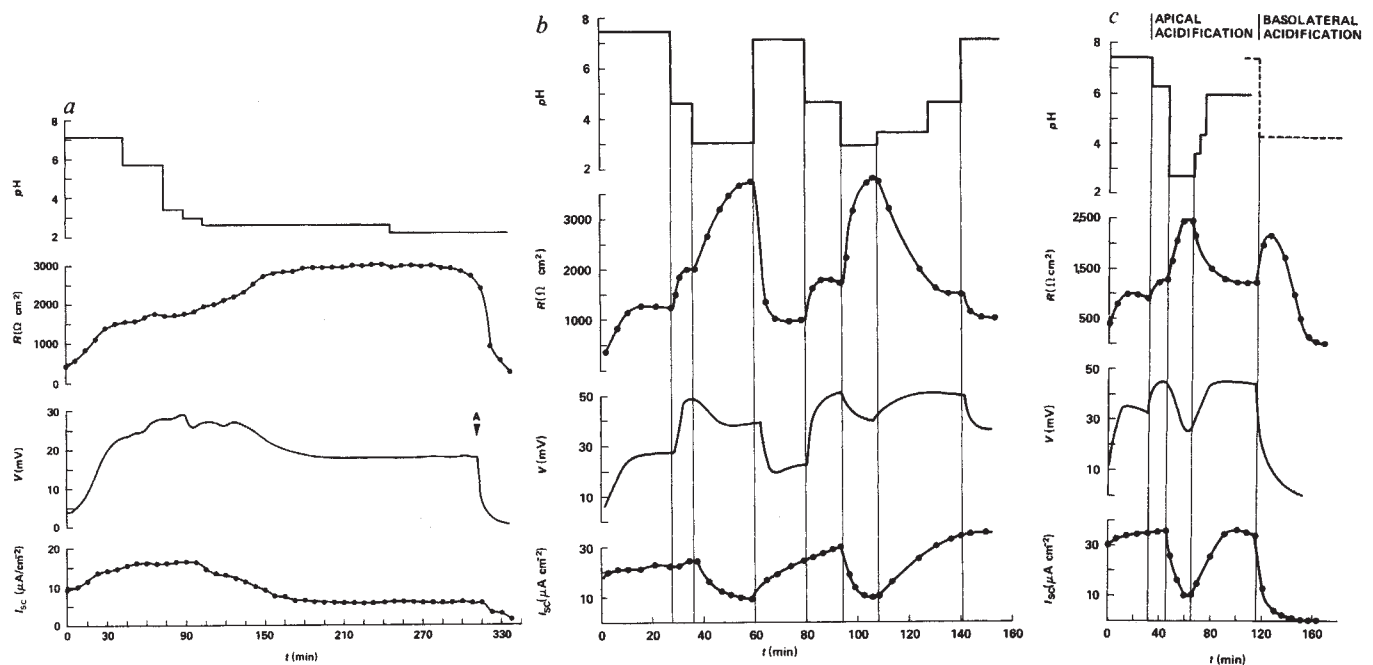


Fig. 1 The electrical response of a chief cell monolayer to mucosal acidification. Top panels: apical solution pH as a function of time; the decrease reflects the addition of acid. Upper middle panels: membrane resistance (R). Lower middle panels: potential difference (V). Bottom panels: short-circuit current (I_{sc}). *a*, With the reduction of apical pH to <2.5 , R increased and remained stable for over 3 h. 4 mM aspirin (A), added at the arrow, caused a rapid decay in R , V and I_{sc} . *b*, The reversibility of the changes in R , V and I_{sc} in response to apical acidification. In occasional monolayers, if the apical pH was decreased <3 , I_{sc} did not fully return to control values on apical neutralization. *c*, A typical response to apical acidification, followed by return of apical pH to 6. After recovery of the electrical parameters (not to control values because the pH was returned to only 6), the serosal solution was acidified to pH 4.4, causing a rapid fall in I_{sc} and V . The transient increase in R following basolateral acidification occurred in 9 of 11 experiments. *d*, Period 1: ouabain (1 mM) caused a drop in both I_{sc} and V , although not to zero. Period 2: acidification of the apical solution to pH 3 produced the same biphasic change in R found in untreated monolayers, although V and I_{sc} decreased to essentially zero in these ouabain-treated monolayers. After neutralization, R returned to near its initial value. Period 3: acidification to pH 2.7 produced a similar pattern to that observed in period 2, but when the pH was reduced to 2.4, damage (decay in R) occurred and a V and I_{sc} became detectable. Neutralization (abolition of the pH gradient) eliminated V and I_{sc} . Period 4: during re-acidification to pH 2.6, V and I_{sc} remained zero, indicating some restoration of monolayer impermeability to H^+ . At lower pH values, R deteriorated and a V and I_{sc} were recorded; apical neutralization again eliminated this V and I_{sc} , indicating a proton current across the damaged monolayer in response to a large H^+ gradient.

Methods: Cells from canine fundic mucosa were dispersed by sequential treatment with collagenase and EDTA and separated by size using counterflow centrifugation in an elutriator rotor^{3,13}. Cells from fractions 4 and 5 in the previously described separation protocol³ were plated in complete culture medium onto collagen gels at $0.7\text{--}1.0 \times 10^6$ cells cm^{-2} . The monolayers that formed after 48 h contained 96% chief cells, as demonstrated by immunofluorescence with anti-pepsinogen antibody³. Mucous cells staining with periodic acid Schiff accounted for $<2\%$ of the cells in the monolayer, although non-adherent PAS⁺ cells were present in the culture medium. After confluency had been obtained, gels were freed from the culture wells and allowed to float in the medium. The monolayers were mounted 12–25 h later in Ussing chambers between plastic inserts with an aperture of 0.66 cm^2 ; a nylon mesh was used for mechanical support and silicon grease was applied to minimize edge damage. V was measured continuously and R calculated from the deflections in V produced by $10\text{-}\mu\text{A}$ pulses passed across the monolayer. I_{sc} was measured using a voltage clamp or calculated from R and V . Bathing solutions were acidified with 1 M HCl. Polarity of the cells in these monolayers was evident by electrical and morphological criteria; monolayers maintained a high apical negative V (range 21–34 mV) and displayed apical junctional complexes and microvilli³. The bathing solution was HCO_3^- -free Hank's salt solution (pH 7.2) bubbled with air.

a potential difference across apically acidified, ouabain-treated monolayers indicates that intact monolayers are practically impermeable to protons. However, when the apical pH was reduced to a point at which R began to decay, a potential difference and a short-circuit current developed across the now damaged monolayer (Fig. 1*d*). We conclude that this I_{sc} reflected H^+ flux, as neutralization of the apical solution eliminated the current (Fig. 1*d*).

The re-establishment of apical to basolateral polarity by dispersed chief cells in monolayer culture³ is evident from the resistance of the apical, but not basolateral, surface to H^+ back-diffusion. This apical specialization is of advantage to chief cells because in gastric glands their apical surfaces are exposed to very high acid and pepsin concentrations. Whether the apical surface of other mucosal cells possess such impermeability to H^+ is unknown. Mucosal defence is probably a multifactorial

process and different cells may use different mechanisms to resist injury. For example, surface epithelial cells secrete bicarbonate and mucus, and mucus may provide an 'unstirred layer' at the cell surface to stabilize a pH gradient generated by HCO_3^- secretion^{10,11}. Furthermore, such cells replicate rapidly and are sloughed off from regions of mucosal injury, with epithelial integrity quickly being restored by migration of adjacent cells¹². As a result of these other mechanisms, surface epithelial cells may be less dependent on an H^+ -impermeable apical surface than are chief cells, which replicate slowly, lack a surface mucus coat and appear not to secrete bicarbonate. (Neither transmission³ nor scanning electron microscopy reveal a mucus coat on the apical surface of chief cell monolayers. Although cells

staining with periodic acid Schiff (PAS^+) constitute 40% of the plated cells, only 1% of monolayer cells are PAS^+ .)

The study of the mechanisms underlying gastric mucosal defence has been hampered by the cellular heterogeneity of the tissue, which makes intact mucosa a complex model. As chief cell monolayers in primary culture display barrier function, this system provides a simple model for studying this element of mucosal defence. Our present findings indicate that chief cells in culture re-form an apical surface virtually impermeable to protons; we suggest that this mechanism underlies chief cell resistance to acid-peptic injury.

This work was supported by NIH grant AM-30444 and a grant from the Stiftung Volkswagenwerk.

Received 3 September; accepted 21 September 1984.

1. Davenport, H. W., Warner, H. A. & Code, C. F. *Gastroenterology* **47**, 142-152 (1964).
2. Davenport, H. W. in *Progress in Gastroenterology* Vol. 2 (ed. Glass, G. B. J.) 42-56 (Grune & Stratton, New York, 1970).
3. Ayalon, A., Sanders, M. J., Thomas, L. P., Amiran, D. A. & Soll, A. H. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7009-7013 (1982).
4. Sanders, M. J., Amiran, D. A., Ayalon, A. & Soll, A. H. *Am. J. Physiol.* **245**, G641-G646 (1983).
5. Powell, D. W. *Am. J. Physiol.* **241**, G275-G288 (1981).
6. Machen, T. E., Silen, B. & Forte, J. *Am. J. Physiol.* **234**, E228-E235 (1978).

7. Wright, E. M. & Diamond, J. M. *Biochim. biophys. Acta* **163**, 57-74 (1978).
8. Mandel, L. J. *Am. J. Physiol.* **235**, C35-C48 (1978).
9. Schiessel, R., Matthews, J., Barzilai, A., Merhav, A. & Silen, W. *Nature* **283**, 671-673 (1980).
10. Flemstrom, G., Heylings, J. R. & Garner, A. *Am. J. Physiol.* **242**, G100-G110 (1982).
11. Grossman, M. I. in *Basic Mechanisms of Gastrointestinal Mucosal Cell Injury and Protection* (ed. Harmon, J.) 323-326 (Williams & Wilkins, New York, 1981).
12. Rutten, M. J. & Ito, S. *Am. J. Physiol.* **244**, G171-G182 (1983).
13. Soll, A. H. *J. clin. Invest.* **61**, 370-380 (1978).

Calcitonin gene-related peptide is a potent vasodilator

S. D. Brain*, T. J. Williams*, J. R. Tippins†, H. R. Morris† & I. MacIntyre‡

Department of Pharmacology, Institute of Basic Medical Sciences, Royal College of Surgeons of England, Lincoln's Inn Fields, London WC2A 3PN, UK

† Department of Biochemistry, Imperial College, London SW7 2AZ, UK

‡ Department of Chemical Pathology, Royal Postgraduate Medical School, DuCane Road, London W12 0HS, UK

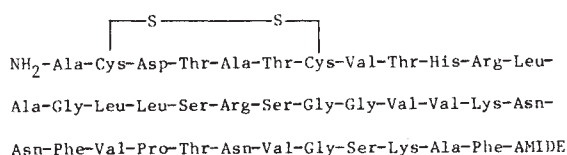
A novel peptide, calcitonin gene-related peptide (CGRP), has been predicted to result from alternative processing of the primary RNA transcript of the calcitonin gene in the rat^{1,2}. Several lines of evidence suggest that CGRP is a transmitter in the central and peripheral nervous system²⁻⁴. Human CGRP has been isolated and characterized⁵, and shown to have potent effects on the heart⁶. The observations presented here indicate that human and rat CGRP also have potent effects on blood vessels. Intradermal injection of CGRP in femtomole doses induces microvascular dilatation resulting in increased blood flow, which we have detected in the rabbit by using a ¹³³Xe clearance technique^{7,8}. In human skin, CGRP induces persistent local reddening. Microscopic observation of the hamster cheek pouch⁹ *in vivo* revealed that topical application of CGRP induces dilatation of arterioles. Furthermore, CGRP relaxes strips of rat aorta *in vitro* by an endothelial cell-dependent mechanism. Therefore, we suggest that local extravascular release of CGRP may be involved in the physiological control of blood flow and that circulating CGRP may contribute to hyperaemia in certain pathological conditions.

The primary RNA transcript of the calcitonin gene can be processed to give alternatively spliced messenger RNA (mRNA) products: one of these encodes the precursor of the 32-amino acid thyroid hormone, calcitonin, and another the precursor of a putative neuropeptide, the 37-amino acid CGRP^{1,2}. The primary structure of rat CGRP has been predicted by analysis of the calcitonin gene coding sequence¹. Immunoreactivity to the C-terminus of the predicted CGRP peptide has been detected in discrete areas of the central and peripheral nervous system of the rat². Rat trigeminal ganglionic cells release immunoreactive CGRP *in vitro*, which is increased by a high potassium

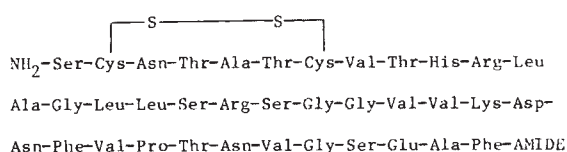
concentration in the medium⁴. Intracerebroventricular injection of CGRP stimulates sympathetic outflow *in vivo* as evidenced by increased levels of circulating noradrenaline, tachycardia and elevated blood pressure³. Thus, CGRP seems to be a neuropeptide. However, in these experiments, control injections of CGRP given intravenously (i.v.) were shown to induce tachycardia accompanied by lowered blood pressure, suggesting peripheral vasodilatation³. Our experiments offer an explanation for these control observations.

Human CGRP was extracted from medullary thyroid carcinoma tissue and purified to homogeneity, as described previously⁵. Chemically synthesized rat CGRP was obtained as a commercial preparation (Bachem). The structures of human CGRP⁵ and rat CGRP¹ are shown below.

Human CGRP



Rat CGRP



Both peptides were tested for vasodilator activity using a multiple-site ¹³³Xe clearance technique in rabbit skin^{7,8}. The peptides were mixed with a solution of ¹³³Xe and injected rapidly in 100-μl volumes into marked sites on the clipped dorsal skin of New Zealand White (NZW) rabbits. Intradermal injections were given in random-block order according to a predetermined balanced site pattern, with six replicates for each of seven treatments per animal. CGRP doses were tested against standard doses of a known potent vasodilator, prostaglandin E₂ (PGE₂), and saline controls. Animals were killed after 15 min, the dorsal skin excised and injected sites removed with a 17-mm diameter punch. Skin samples and samples of injection fluids were measured for radioactivity in an automatic γ-counter. Changes in blood flow induced by the test substances were determined as described in Fig. 1 legend. Both human and rat CGRP increased local blood flow with potencies similar to that of the highly active PGE₂. Blood flow was significantly increased with

* Present address: Section of Vascular Biology, MRC Clinical Research Centre, Watford Road, Harrow, Middlesex HA1 3UJ, UK.

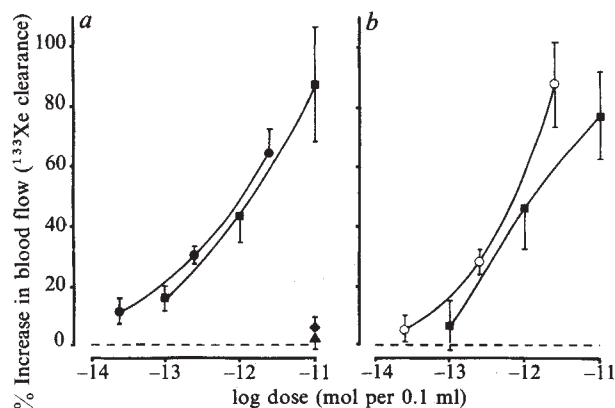


Fig. 1 Effect of human and rat CGRP on blood flow in rabbit skin. *a*, A comparison of increased blood flow induced by human CGRP (●) with that induced by PGE₂ (■) in seven rabbits. Both CGRP and PGE₂ increased blood flow in a dose-dependent manner, compared with the control (phosphate-buffered saline (PBS), broken line). Human CGRP at 25 fmol induced a significant change in blood flow ($P < 0.02$, paired Student's *t*-test). Human calcitonin (10 pmol, ◆) and katalcacin (10 pmol, ▲) were also tested in four rabbits but neither had any detectable effect on blood flow. *b*, A comparison of increased blood flow induced by rat CGRP (○) and PGE₂ (■) in five rabbits. Rat CGRP at 250 fmol induced a significant change in blood flow ($P < 0.01$, paired Student's *t*-test). **Methods:** Experiments were carried out at room temperature (22–25 °C) with NZW rabbits (3.5–4 kg) anaesthetized with Saffan (0.5 ml per kg). Agents mixed with ^{133}Xe (5–10 μCi per injection) were injected intradermally in 100- μl volumes into the clipped back skin in random-block order according to a balanced site pattern. The agents tested were CGRP (0.025–2.5 pmol), PGE₂ (0.1–10 pmol) for comparison, and PBS as a control; six replicates of each dose were tested in each rabbit. After 15 min, the animal was killed with an i.v. barbiturate overdose (Espiril), the back skin was removed and a 17-mm diameter punch was used to excise injection areas. Samples of skin and injection fluid, under paraffin oil in sealed tubes, were counted immediately in an automatic γ -counter. The change in local blood flow was calculated as the per cent increase above PBS controls from the equation: $100 (\ln^{133}\text{Xe count of saline-injected skin} - \ln^{133}\text{Xe count of agent-injected skin}) / (\ln^{133}\text{Xe count of 0.1 ml injection fluid} - \ln^{133}\text{Xe count of saline-injected skin})$. Results are expressed as mean $\pm$ s.e.m. for *n* rabbits.

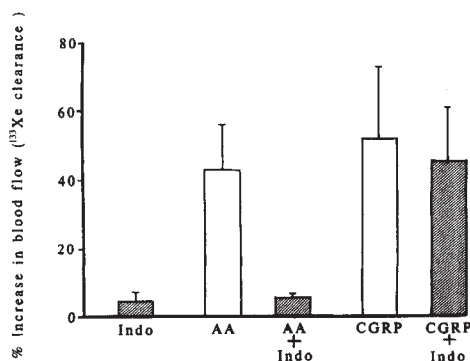


Fig. 2 Effect of the prostaglandin synthesis inhibitor, indomethacin, on increased blood flow induced by human CGRP and arachidonic acid (AA) in rabbit skin. Indomethacin (Indo, 2.8 nmol) or PBS was injected intradermally 15 min before the test compounds. ^{133}Xe was mixed with the test compounds, and human CGRP (2.5 pmol) or arachidonic acid (3.3 nmol) was then injected into the same sites with a further dose of indomethacin (2.8 nmol) as required. We determined the responses over a 15-min period as described for Fig. 1. Indomethacin had no effect on the response to CGRP, but caused a large inhibition of increased blood flow induced by arachidonic acid ($P < 0.02$, paired Student's *t*-test). Indomethacin alone had no significant effect on blood flow. Results are expressed as mean $\pm$ s.e.m. for four rabbits.

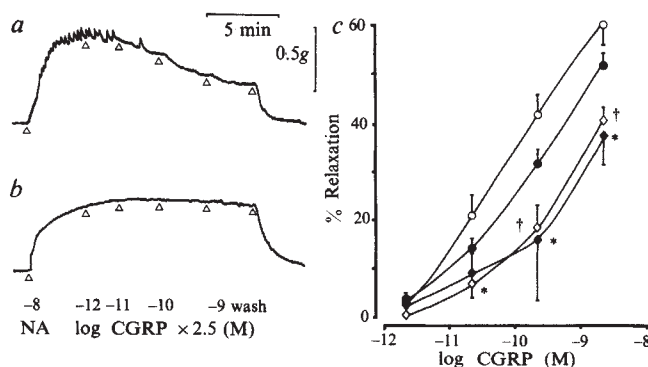


Fig. 3 Response of rat aortic rings to human and rat CGRP. Rat thoracic aorta was cleared of fat and connective tissue and cut transversely into rings 3 mm wide. Each ring was cut to form a strip of circular muscle, taking care not to damage the endothelium¹³. The strip was then held in a 10-ml organ bath in oxygenated Krebs' solution containing EDTA (30 μM)¹³. The tissue was suspended from an isometric transducer (Lectromed UFI) under 1g tension and was allowed to equilibrate for 90 min before the experiment. *a*, Responses of aorta with undamaged endothelium; *b*, responses of aorta after the endothelium has been gently rubbed with a microspatula before mounting in the organ bath. The aortae were induced to contract by adding noradrenaline (NA, 10^{-8} M) to the bath, which sometimes induced some spontaneous activity in preparations with intact endothelium. At the point of maximum contraction, a control response was obtained by adding cumulative concentrations of ACh (10^{-8} – 10^{-6} M) to the bath (results not shown). Only when the endothelium was intact did ACh produce a dose-related relaxation of the preparation; ~50% relaxation was achieved with 10^{-7} M and 100% relaxation with 10^{-6} M ACh. No relaxation was induced by ACh after removal of the endothelium. Human CGRP, added in cumulative concentrations, produced a dose-related relaxation of the aorta with an intact endothelium; the aorta in which the endothelium had been destroyed was not affected. Responses to CGRP took longer to reach equilibrium than those to ACh (~3 min for CGRP compared with 1 min for ACh). *c*, Dose-response curves for human and rat CGRP on NA-contracted rat aorta to show the effect of indomethacin. Indomethacin (2.8×10^{-6} M final concentration) was added to the bath 15 min before the addition of NA and the cumulative doses of CGRP. ●, Relaxation induced by human CGRP; ○, human CGRP in the presence of indomethacin; ◇, rat CGRP; ◇, rat CGRP in the presence of indomethacin. Each point is the mean $\pm$ s.e.m. of 4–7 observations. Indomethacin did not affect relaxation induced by ACh, but partially inhibited that induced by CGRP (* $P < 0.05$; † $P < 0.005$, unpaired Student's *t*-test).

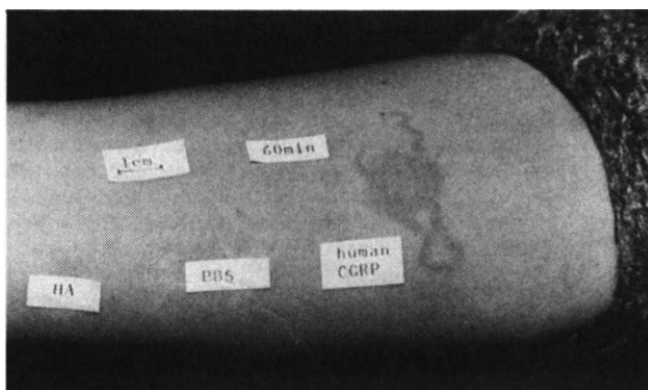


Fig. 4 Effect of intradermal injection of CGRP into human skin. Synthetic human CGRP (15 pmol; Sandoz) was prepared in 50 μl sterile PBS. Histamine (HA, 500 pmol per 50 μl) was injected with PBS as a control. The figure shows the appearance of the injected arm 60 min after injection, at which time the response to histamine had faded. The dilator response to CGRP is clearly visible as local reddening, spreading slowly, with pseudopodia. This response persisted for 5–6 h.

a 25-fmol dose of human CGRP ($P < 0.02$, Student's *t*-test). No changes in blood flow were detected in response to either human calcitonin or katecalcin, a 21-amino acid peptide derived from the same precursor¹⁰. Human or rat CGRP, at doses up to 2.5 pmol, did not cause microvascular plasma protein leakage, as measured by the local accumulation of ¹²⁵I-albumin injected i.v.⁸. This shows that the response to CGRP does not involve the release of histamine, which increases vascular permeability.

To test whether the dilator responses to CGRP in rabbit skin were caused by release of endogenous prostaglandins (both PGE₂ and prostaglandin I₂ (PGI₂)) are very potent dilators in the rabbit⁸), human CGRP was injected into skin sites preinjected with indomethacin, which inhibits prostaglandin synthesis. Arachidonic acid was used as a positive control; intradermally injected arachidonic acid is converted to a dilator prostaglandin in this system^{8,11}. Figure 2 shows that increased blood flow induced by human CGRP in the skin was unaffected by indomethacin, whereas responses to arachidonic acid were inhibited. In this respect the effect of CGRP is similar to that of another putative neuropeptide, vasoactive intestinal polypeptide (VIP)¹².

We also investigated the possibility of a secondary release of other vasodilators in the skin by CGRP. The vasodilator responses to CGRP were found to be unaffected by the acetylcholine (ACh) antagonist, hyoscine (0.1 nmol per site) and the β -adrenergic antagonist, propranolol (0.1–1 nmol per site). Moreover, as reported previously¹², ACh, isoprenaline, purines and substance P are weak vasodilators in rabbit skin: CGRP was at least 1,000 times more potent than ACh, ATP, ADP, adenosine, 5-hydroxytryptamine and substance P, and 10–100 times more potent than the synthetic β -adrenergic stimulant isoprenaline.

Direct microscopic observation of the hamster cheek pouch preparation *in vivo*⁹ confirmed that increased ¹³³Xe clearance induced by CGRP was caused by dilatation of the major resistance vessels, the arterioles. The microvascular bed was superfused with modified Krebs saline solution containing indomethacin (2.8×10^{-6} M). Topical application of CGRP (2.5 pmol per 10 μ l) to preparations with a high arteriolar tone (spontaneous, or induced with a superfusion of noradrenaline (1.5×10^{-7} M) in two experiments) caused arteriolar dilatation from an initial diameter of 12.8 ± 3.5 μ m to 39.3 ± 6.3 μ m (mean $\pm$ s.e.m., $n = 4$ animals; $P < 0.02$, paired Student's *t*-test). The arteriolar dilatation in response to CGRP was maximal 2 min after application and persisted for at least 5 min.

Immunoreactive CGRP has been detected in extracts of rat aortae (20 pmol per g in a pooled extract of 12 aortae; S. Wimalawansa, unpublished observation). We have discovered that CGRP has a relaxant effect on artery preparations *in vitro* by an endothelial cell-dependent mechanism. It has been reported previously that some substances induce vasodilatation by stimulating the release, from vascular endothelial cells, of an unknown secondary mediator which causes smooth muscle relaxation¹³. CGRP was tested in this manner using 3 mm-wide transverse strips of rat thoracic aorta *in vitro*. Noradrenaline (10^{-8} M) was added to the bathing medium to induce tone; human CGRP was then added in increasing concentrations. A progressive relaxation of the tissue was seen (Fig. 3a), which was slower than that induced by ACh (see Fig. 3 legend). Relaxation was not seen in aortic strips from which the endothelial cells had been removed, as shown in Fig. 3b. A

comparison of endothelium-dependent relaxation induced by human and rat CGRP is shown in Fig. 3c; each produced relaxation at concentrations as low as 2.5×10^{-11} M. Indomethacin pretreatment partially suppressed the relaxation induced by both CGRP preparations, but not that induced by ACh (see Fig. 3 legend). Thus, prostaglandin generation may be responsible for a part of the response to CGRP. A contribution by prostaglandins should be considered in future studies of CGRP effects, although there was no evidence of such a contribution in the rabbit skin experiments.

The potent vasodilator effects of CGRP were also observed in human volunteers (Fig. 4). Synthetic human CGRP (Sandoz, Basle, Switzerland) was used, and eighteen doses injected intradermally into the volar surface of the forearm of three non-atopic subjects (S.D.B., J.R.T. and T.J.W.). CGRP (15 fmol–15 pmol per 50 μ l) induced vasodilation that was seen as local reddening within 1 min after injection. The reddening became more intense, while spreading slowly. The response lasted for 1 h with a 15-fmol dose and 5–6 h with a 15-pmol dose. At the higher dose, local reddening spread in streaks, suggesting lateral dispersion via lymphatic vessels¹⁴. A transient patchy 'flare' surrounding the local reddening was sometimes observed, but there were no obvious weals over this dose range. A higher dose of CGRP, 250 pmol, induced a weal and flare response comparable with that induced by histamine at 500 pmol. However, local reddening induced by histamine was less intense and faded with the weal and flare within an hour, whereas local reddening induced by the high dose of CGRP persisted for 10–12 h.

These experiments demonstrate that CGRP is a potent vasodilator. Relaxation of the aorta *in vitro* seems to depend on endothelial cells, which may be relevant to the responses of major vessels and microvessels to circulating CGRP. Circulating CGRP could be involved in hyperaemia such as the flushing sometimes found in patients with medullary thyroid carcinoma¹⁵, where CGRP has been detected in the plasma⁵. Whether arteriolar dilatation induced by local extravascular CGRP (as in the skin and hamster cheek pouch experiments) is endothelial cell-dependent remains open to question for this and other dilators such as ACh. It is possible that CGRP receptors are present on the abluminal surface of endothelial cells or, alternatively, arteriolar smooth muscle cells may have their own CGRP receptors. Immunoreactive CGRP has been observed in nerve fibres associated with blood vessels², which raises the possibility that CGRP is involved in neurogenic inflammation¹⁶. CGRP may be the endogenous mediator of the skin flare response, as an alternative to substance P which has been proposed to be released from nerve terminals by antidromic impulses passing down branches of afferent nerves¹⁷.

We know of no other reports of vasodilators producing significant effects at low femtomole doses. The very high dilator potency of CGRP, shown here in four unrelated species, implies a function for this novel peptide. We consider it worthwhile to investigate the possibility that CGRP has a role in the physiological control of vascular tone and blood flow.

We thank Dr S. Girgis for extraction of human CGRP, Dr M. Panico and Mr T. Etienne for purification and quantitation of CGRP. The human skin experiments were carried out with permission of the Ethical Committee of the Royal College of Surgeons of England. We thank the Wellcome Trust, the British Heart Foundation and the MRC for financial support.

Received 30 June; accepted 28 September 1984.

- Amara, S. G., Jonas, V., Rosenfeld, M. G., Ong, E. S. & Evans, R. M. *Nature* **298**, 240–244 (1982).
- Rosenfeld, M. G. *et al.* *Nature* **304**, 129–135 (1983).
- Fisher, L. A. *et al.* *Nature* **305**, 534–536 (1983).
- Mason, R. T. *et al.* *Nature* **308**, 653–655 (1984).
- Morris, H. R. *et al.* *Nature* **308**, 746–748 (1984).
- Etienne, T. *et al.* *J. Physiol., Lond.* **351**, 48P (1984).
- Williams, T. J. *J. Physiol., Lond.* **254**, 4–5P (1976).

- Williams, T. J. *Br. J. Pharmac.* **65**, 517–524 (1979).
- Duling, B. R. *Microvasc. Res.* **5**, 423–429 (1973).
- Hillyard, C. J. *et al.* *Lancet* **i**, 846–848 (1983).
- Williams, T. J. & Peck, M. J. *Nature* **270**, 530–532 (1977).
- Williams, T. J. *Br. J. Pharmac.* **77**, 505–509 (1982).
- Furchgott, R. F. & Zawadzki, J. V. *Nature* **288**, 373–376 (1980).
- Solomon, L. M., Juhlin, L. & Kirschenbaum, M. B. *J. invest. Derm.* **51**, 280–282 (1968).
- Cohen, S. L., MacIntyre, I., Grahame-Smith, D. & Walker, J. G. *Lancet* **ii**, 1172–1174 (1973).
- Jancso, G., Kiraly, E. & Jancso-Gabor, A. *Nature* **270**, 741–743 (1977).
- Lembeck, F. & Holzer, P. *Naunyn-Schmiedeberg's Archs Pharmac.* **310**, 175–183 (1979).

Isolation and structure of a novel C-terminally amidated opioid peptide, amidorphin, from bovine adrenal medulla

B. R. Seizinger*, D. C. Liebisch*, C. Gramsch*, A. Herz*, E. Weber†, C. J. Evans†, F. S. Esch‡ & P. Böhlen‡

* Department of Neuropharmacology, Max-Planck-Institut für Psychiatrie, Am Klopferspitz 18A, 8033 Martinsried, FRG
† Nancy Pritzker Laboratory of Behavioral Neurochemistry, Stanford University School of Medicine, Stanford, California 94305, USA
‡ Laboratories for Neuroendocrinology, The Salk Institute for Biological Studies, La Jolla, California 92037, USA

Biologically active peptide hormones and neurotransmitters have been shown to be enzymatically liberated from larger, inactive precursor molecules by tissue-specific post-translational processing, particularly at the typical cleavage signals of paired basic residues^{1,2}. Subsequent N-terminal or C-terminal modifications may be of importance in regulating the biological activities of these peptides (for review see ref. 3). C-terminal α -amidation is considered to be essential for the biological function of several non-opioid peptides^{4,5}. Here we present the isolation and structure of a novel C-terminally amidated opioid peptide, amidorphin, from bovine adrenal medulla. Amidorphin and the recently isolated octapeptide metorphamide⁶ (adrenorphin⁷) are the only endogenous opioid peptides in mammals known to possess a C-terminal amide group. The amino acid sequence of amidorphin corresponds to the sequence 104–129 of bovine proenkephalin A (refs 8, 9). Very high concentrations of amidorphin were detected in bovine adrenal medulla and in a further endocrinological system, the hypothalamic–neurohypophyseal axis. Amidorphin may therefore be considered to be a major gene product of the opioid peptide precursor proenkephalin A in these endocrine tissues.

The sequence containing residues 104–130 of bovine proenkephalin A is flanked by pairs of basic lysine–arginine residues^{8–10}, which are considered to be typical cleavage signals for the enzymatic processing of peptide precursors^{1,2}. Processing at these cleavage sites should liberate an opioid peptide with 27 amino acid residues which contains the Met-enkephalin sequence at its N-terminus and a glycine residue at its C-terminus. It has been shown for several non-opioid peptides that after enzymatic processing at a pair of basic residues, a C-terminal glycine residue may serve as a nitrogen donor for the amidation of the preceding amino acid residue¹¹. This conversion is considered to be mediated by a specific amidation enzyme^{11,12}. Thus, an opioid peptide with 26 amino acid residues, which corresponds to the sequence 104–129 of bovine proenkephalin A and carries a C-terminal amide group, may occur naturally. To investigate this hypothesis, the peptide was synthesized by Peninsula Laboratories and a radioimmunoassay (RIA) using a specific antiserum against the synthetic peptide was developed (see Fig. 1b legend). Using this RIA as monitoring system, the peptide, which has been named amidorphin (amidated endorphin-like peptide), was isolated from bovine adrenal medulla. It was purified by gel filtration, affinity chromatography with the monoclonal antibody 3-E7 and finally by three different HPLC steps (see Fig. 1 legend).

Figure 1 shows the absorbance and immunoreactivities of a typical third (final) HPLC purification step. The absorbance peak with the retention time of 84 min eluted in the position of synthetic amidorphin (Fig. 1a, hatched peak). It was recognized by the amidorphin antiserum, which is specific for the C-terminus of the peptide (Fig. 1b, black bar), and by the monoclonal antibody, specific for the N-terminal Tyr-Gly-Gly-Phe sequence, common to all mammalian endogenous opioid peptides (Fig. 1c, black bar).

Sequence analysis yielded unambiguous results for the first 24 amino acid residues (Table 1a). Amino acid composition

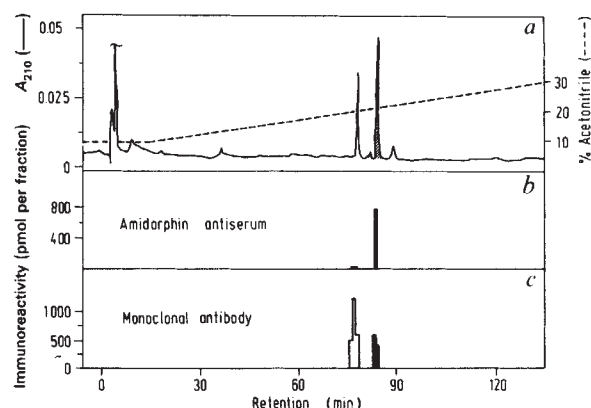


Fig. 1 Final HPLC purification step of amidorphin from bovine adrenal medulla. *a*, Absorbance at 210 nm. Hatched peak indicates absorbance of purified endogenous amidorphin. *b*, Amidorphin immunoreactivity profile, determined with antiserum 'Ekki' directed against synthetic amidorphin. Black bar represents amidorphin-immunoreactivity under the hatched peak. *c*, Immunoreactivity profile, determined with the monoclonal antibody 3-E7. This antibody is specific for the N-terminal sequence Tyr-Gly-Gly-Phe (ref. 17), common to all mammalian endogenous opioid peptides known. The RIA was performed as described elsewhere¹⁷. Black bar represents 3-E7-immunoreactivity under the hatched peak. **Methods:** The EKKi antiserum was produced in New Zealand rabbits as previously described¹⁸ and used in a final dilution of 1:2,500 in a RIA with ¹²⁵I-labelled synthetic amidorphin as tracer. The RIA procedure was as described elsewhere¹⁸, with the modification that the RIA buffer did not contain Triton X-100. The highly specific amidorphin RIA required the C-terminal amide group as a recognition site. On a molar basis, synthetic amidorphin(20–26)-NH₂ cross-reacted by 100% and synthetic amidorphin(20–26)-OH (without the C-terminal amide group) by 0.01%. (Peptides were synthesized as previously described⁶.) Peptide F cross-reacted by 0.4%; the following peptides cross-reacted by <0.01%: metorphamide, Met-enkephalin, Met-enkephalin-NH₂, Met-enkephalin-Arg⁶-Phe⁷, Met-enkephalin-Arg⁶-Phe⁷-NH₂, peptide E, BAM-12P, BAM-22P, Met-enkephalin-Arg⁶-Gly⁷-Leu⁸, β -endorphin, Leu-enkephalin, dynorphin A(1–8), dynorphin A(1–17), α / β -neo-endorphin, α / γ -melanocyte-stimulating hormone. (The amidorphin antiserum will be described in detail elsewhere.) Amidorphin was isolated from fresh bovine adrenal medullas. Within 5 min postmortem, adrenals were placed on ice; within 90 min, they were separated from cortex, sliced and incubated in 5 vol/wt 0.1 M HCl for 10 min at 96°C. Tissues were homogenized, centrifuged and supernatants freed from lipids with 5 vol heptane/vol supernatant. After evaporating, the extracts were redissolved in 35 ml of 0.1 M acetic acid and chromatographed on a Sephadex G-50 fine column (5 × 90 cm), eluting in 0.1 M acetic acid. Fractions (10 ml) were collected at a flow rate of 100 ml h⁻¹ and aliquots determined for immunoreactive amidorphin using a RIA with Ekki. After gel filtration, amidorphin was further purified by affinity chromatography using the monoclonal antibody 3-E7. This antibody recognizes the N-terminal sequence Tyr-Gly-Gly-Phe (ref. 17), common to all endogenous opioid peptides (for coupling procedures see ref. 19). 30 ml gel was packed in a column (1 × 40 cm) and equilibrated in 0.1 M ammonium acetate, pH 7.4. The pooled and lyophilized fractions from gel filtration, containing amidorphin in a single peak in the position of synthetic amidorphin (not shown) were redissolved in 0.1 M ammonium acetate buffer, pH 7.4, centrifuged and supernatants applied to the affinity column at 4°C (flow rate 0.1 ml min⁻¹). Subsequently, the column was washed with 500 ml 0.1 M ammonium acetate and eluted with 40 ml 2.5 M acetic acid containing 1 M ammonium acetate. The eluate was lyophilized, dissolved in water and kept frozen before HPLC. Final purification of amidorphin was achieved by three different HPLC steps: (1) an Altex Ultrasphere ODS column (4.6 × 250 mm, particle size 5 μ m), buffer: 50 mM NaH₂PO₄/methanol (95%/5%) pH 2.7, acetonitrile elution gradient 10–40% in 120 min; (2) an Altex-Ultrapore RPSC column (4.6 × 75 mm), buffer 0.1% trifluoroacetic acid (TFA), acetonitrile elution gradient 10–40% in 120 min; (3) an Altex Ultrasphere ODS column (4.6 × 250 mm; particle size 5 μ m), buffer: 50 mM NaH₂PO₄/methanol (95%/5%) pH 7.0, acetonitrile elution gradient 10–30% in 120 min (see Fig. 1). For all HPLC systems, 1.1-ml (1 min) fractions were collected at a flow rate of 1.1 ml min⁻¹ and aliquots assayed for immunoreactive peptides. In all three HPLC systems, total amidorphin immunoreactivity eluted in a single peak in position of synthetic amidorphin. The amidorphin-containing fractions from the third HPLC purification step were desalted with a HPLC Altex Ultrasphere ODS column (6.6 × 75 mm; particle size 3 μ m; buffer: 0.1% TFA; acetonitrile elution gradient 10–40% in 30 min) before sequence analysis, amino acid composition analysis and enzymatic cleavage. Results from four separations with ~80 g tissue each were consistent, each yielding 500–800 pmol of pure amidorphin (as determined by RIA).

analysis showed the presence of 26 amino acid residues (Table 1b), 24 of which were identified by automated Edman degradation. The two additional amino acid residues, valine and leucine, could thus be assigned to the C-terminus. The primary structure of amidorphin corresponds to the first 26 amino acid residues of peptide F (ref. 13), which is contained in the sequence 104–129 of bovine proenkephalin A^{8,9}.

The C-terminal amide group of amidorphin was confirmed by carboxypeptidase A digestion experiments and by the immunological characteristics of the peptide. Endogenous amidorphin was completely stable to carboxypeptidase A digestion over a time period of at least 5 h (Fig. 2) and exhibited full cross-reactivity with synthetic amidorphin in the amidorphin RIA that required the C-terminal amide residue as a recognition site (see Fig. 1b legend). Thus, we conclude that amidorphin was amidated at the C-terminus and that the complete primary structure of the peptide is Tyr-Gly-Gly-Phe-Met-Lys-Lys-Met-Asp-Glu-Leu-Tyr-Pro-Leu-Glu-Val-Glu-Glu-Ala-Asn-Gly-Gly-Glu-Val-Leu-NH₂.

Amidorphin could be detected in extremely high concentrations in bovine adrenal medulla and the hypothalamic-neurohypophyseal axis (Table 2). In the adrenal medulla, total amidorphin immunoreactivity consisted almost exclusively of authentic amidorphin, which was determined by a combined use of amidorphin RIA, monoclonal 3-E7 RIA and co-elution with synthetic amidorphin on gel filtration and HPLC. In addition

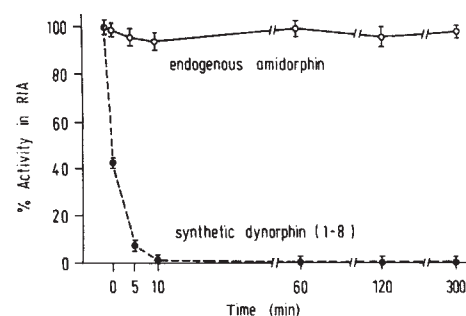


Fig. 2 Effect of carboxypeptidase A digestion on endogenous amidorphin from bovine adrenal medulla (closed circle) and synthetic dynorphin(1–8) (open circles). Endogenous amidorphin was stable to carboxypeptidase A digestion for at least 5 h whereas dynorphin(1–8), which was added together with amidorphin to the enzymatic solution as a non-amidated control peptide, was completely digested within 10 min.

Methods: A mixture of endogenous amidorphin (~130 pmol) and synthetic dynorphin(1–8) (100 pmol) was subjected to carboxypeptidase A digestion according to the method previously described⁶. Briefly, the peptides were dissolved in 3 ml of 0.02 M Tris-HCl/0.2 M NaCl/0.1% Triton X-100, pH 7.5. A suspension (10 µl) of carboxypeptidase A (20 mg ml⁻¹) (Sigma, type 1) containing 200 µg of enzyme was dissolved in 1 ml of 10% LiCl. 5 µl of the freshly prepared solution containing 1 µg of enzyme was added to the peptide mixture and the sample was incubated at 37 °C. Six 40-µl aliquots were withdrawn at, respectively, 1, 5, 10, 60, 120 and 300 min; to terminate the enzyme reaction, each sample was incubated for 10 min at 96 °C. Finally, the samples were assayed in triplicate by RIA for amidorphin and dynorphin (1–8) (dynorphin(1–8) was used as a control substance, because for this peptide a RIA was available that is specific for the C-terminus of dynorphin(1–8)²⁰). Zero-time samples were withdrawn before addition of enzyme to determine the amount of peptide initially present in the sample. Separate samples of peptide mixture either with or without enzyme solution after inactivation at 96 °C for 10 min were incubated for 300 min and assayed for immunoreactive peptides, in order to exclude the possibility of non-specific degradation of peptides during incubation.

Table 1 Gas-phase amino acid sequence analysis (a) and amino acid composition analysis (b) of bovine adrenal medulla amidorphin

a	Cycle no. (N)	>PhNCS amino acid	Yield (pmol)	Cycle no. (N)	>PhNCS amino acid	Yield (pmol)
	1	Tyr	165	15	Glu	40
	2	Gly	52	16	Val	26
	3	Gly	83	17	Glu	38
	4	Phe	66	18	Glu	32
	5	Met	168	19	Glu	49
	6	Lys	16	20	Ala	43
	7	Lys	40	21	Asn	36
	8	Met	96	22	Gly	35
	9	Asp	48	23	Gly	47
	10	Glu	45	24	Glu	20
	11	Leu	60	25	—	—
	12	Tyr	85	26	—	—
	13	Pro	41	27	—	—
	14	Leu	61			

b	Amino acid	No. of residues per mol	Nearest integer	Theoretical values
	Asx	2.19	2	2
	Thr	0	0	0
	Ser	0.36	0	0
	Glx	5.97	6	6
	Gly	4.18	4	4
	Ala	1.16	1	1
	Val	1.87	2	2
	Met	1.65	2	2
	Ile	0	0	0
	Leu	2.98	3	3
	Tyr	1.93	2	2
	Phe	0.83	1	1
	His	0	0	0
	Trp	0	0	0
	Lys	1.75	2	2
	Arg	0	0	0
	Cys	0	0	0
	Pro	1.19	1	1
	Total no. of residues		26	26

a, Gas-phase amino acid sequencing data of bovine amidorphin. 340 pmol (by amino acid analysis) amidorphin was loaded on an Applied Biosystems 470A protein sequencer. Methanol/HCl was used for conversion of anilinothiazolinone (ATZ)-amino acids, and HPLC on an IBM-cyanopropyl column¹⁵ was used for identification of phenylthiohydantoin (PTH)-amino acids. 150 pmol of polyornithine (relative molecular mass 3,000–15,000, Sigma) were added to the sample to serve as an internal standard during sequencing. b, Amino acid composition analysis of bovine amidorphin. Aliquots (22 pmol) of peptide were hydrolysed and analysed for amino acid composition as described previously¹⁶. Values are not corrected for hydrolysis losses and represent the mean of two analyses. Cys was determined as cysteic acid.

Table 2 Regional distribution of amidorphin in bovine adrenal medulla and hypothalamic-neurohypophyseal system

	Total immunoreactivity (pmol per g tissue)	Authentic peptide (pmol per g tissue)
Adrenal medulla	1,376 ± 242	1,172 ± 206
Posterior pituitary	5,210 ± 888	677 ± 115
Hypothalamus	638 ± 73	168 ± 18

The concentrations of authentic amidorphin are compared with the concentrations of total amidorphin immunoreactivity. Values represent mean ± s.e.m., *n* = 5. **Methods:** Fresh bovine tissues from a local slaughterhouse were placed on ice within 5 min of decapitation and extracted within 90 min as described for the isolation procedure. Amidorphin immunoreactivity was determined by RIA (see Fig. 1b legend). The extracts of each area were chromatographed on a Sephadex G-50 superfine column (1.6 × 90 cm), eluting in 15% (v/v) acetic acid. Fractions (3 ml) were collected at a flow rate of 15 ml h⁻¹ and aliquots assayed for immunoreactive amidorphin. In each case amidorphin eluted in a single peak. The peak fractions were treated with H₂O₂ to oxidize amidorphin-related peptides to their respective methionine sulfoxide forms and subjected to a HPLC Altex Ultrasphere ODS column (4.6 × 250 mm; particle size 5 µm), buffer: 50 mM NaH₂PO₄/methanol (95%/5%) pH 7.0; acetonitrile elution gradient 0–30% in 120 min. Fractions (1.1 ml) were collected at a flow rate of 1.1 ml min⁻¹ and aliquots determined with the amidorphin RIA and antibody 3-E7 RIA, respectively (for RIAs see Fig. 1 legend). More than 80% of total amidorphin immunoreactivity from bovine adrenal medulla eluted in the position of synthetic oxidized amidorphin on HPLC. This peptide was also recognized by the RIA with the monoclonal 3-E7 antibody. In contrast, only 25% and 13% of immunoreactive amidorphin from hypothalamus and posterior pituitary, respectively, eluted in the position of synthetic oxidized amidorphin. In these tissues, the amidorphin immunoreactivity was predominantly constituted by peptides that were recognized only by the amidorphin antiserum specific for the C-terminus of the peptide but not by antibody 3-E7, which detects the N-terminus of opioid peptides. Thus, these peptides, which exhibited slightly lower relative molecular masses on gel filtration and shorter retention times on HPLC compared with synthetic oxidized amidorphin, might be C-terminal fragments of amidorphin. (We are now isolating them in order to determine the exact amino acid sequence.)

to authentic amidorphin, hypothalamus and posterior pituitary were found to contain high amounts of putative C-terminal fragments (see Table 2 legend). Thus, in contrast to the adrenal medulla, amidorphin may be further processed in the hypothalamic-neurohypophyseal system, giving rise to C-terminal fragments and, possibly, Met-enkephalin from the N-terminal part of the peptide.

The high concentrations of authentic amidorphin in bovine adrenal medulla and the hypothalamic-neurohypophyseal system suggest that this novel amidated opioid peptide is a major gene product of the opioid peptide precursor proenkephalin A in these endocrine tissues. Note that, for example in bovine adrenal medulla, amidorphin occurs in concentrations almost two orders of magnitude higher than those of metorphamide, the other endogenous opioid peptide known to possess a C-terminal amide group.

The biological significance of the C-terminal amidation of opioid peptides remains unclear. For example, the C-terminal amide group of amidorphin may influence the selectivity of its binding to different opioid receptor subtypes, as has been demonstrated for the C-terminal amide group of metorphamide⁶.

The opioid peptide β -endorphin has been shown to possess a second biologically active (non-opioid) site at the C-terminus¹⁴. Moreover, the presence of the α -amide group is essential for the biological activity³⁻⁵ of several non-opioid peptides possessing C-terminal amide residues, for example corticotropin-releasing factor, gastrin and cholecystokinin. Thus, in addition to a distinctive opioid-like function for amidorphin, it is conceivable that the C-terminal amide group of amidorphin is involved in a non-opioid function of this peptide.

We thank S. Kosin, Dr R. M. Arendt and Dr M. J. Millan for their support.

Received 11 June; accepted 27 September 1984.

- Steiner, D. F., Quinn, P. S., Chan, S. J., Marsh, J. & Tager, H. S. *Ann. N.Y. Acad. Sci.* **343**, 1-16 (1980).
- Li, C. H., Barnafi, L., Chretien, M. & Chung, D. *Nature* **208**, 1093-1094 (1965).
- Mains, R. E., Eipper, B. A., Glembofski, C. C. & Dores, R. M. *Trends Neurosci.* **6**, 229-235 (1983).
- Vale, W., Spiess, J., Rivier, C. & Rivier, J. *Science* **213**, 1394-1397 (1981).
- Rehfeld, J. F. *Am. J. Physiol.* **240**, G255-G266 (1981).
- Weber, E. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 7362-7366 (1983).
- Matsuo, H., Miyata, A. & Mizuno, K. *Nature* **305**, 721-723 (1983).
- Noda, M. *et al. Nature* **95**, 202-206 (1982).
- Gubler, U., Seeburg, P., Hoffman, B. J., Gage, L. P. & Udenfriend, S. *Nature* **295**, 206-208 (1982).

- Comb, M., Seeburg, P. H., Adelman, J., Eiden, L. & Herbert, E. *Nature* **295**, 663-664 (1982).
- Bradbury, A. F., Finnie, M. D. A. & Smyth, D. G. *Nature* **298**, 686-688 (1982).
- Eipper, B. A., Mains, R. E. & Glembofski, C. C. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5144-5148 (1983).
- Kimura, S. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 1681-1685 (1980).
- Schweigerer, L., Bhakdi, S. & Teschemacher, H. *Nature* **296**, 572-574 (1982).
- Hunkapiller, M. & Hood, L. *Meth. Enzym.* **91**, 486-493 (1983).
- Böhlen, P. & Schroeder, R. *Analyt. Biochem.* **26**, 144-152 (1982).
- Gramsch, C., Meo, T., Riethmüller, G. & Herz, A. *J. Neurochem.* **40**, 1220-1226 (1983).
- Höhl, V., Gramsch, C. & Herz, A. in *Radioimmunoassay of Drugs and Hormones in Cardiovascular Medicine* (eds Albertini, A., da Prada, M. & Peskar, B. A.) (Elsevier, Amsterdam, 1979).
- Dandekar, T., Gramsch, C., Houghten, R. A. & Schulz, R. *Neurochem. Int.* (in the press).
- Weber, E., Evans, Ch. J. & Barchas, J. D. *Nature* **299**, 77-79 (1982).

Protein kinase C regulation of the receptor-coupled calcium signal in histamine-secreting rat basophilic leukaemia cells

Ronit Sagi-Eisenberg, Harry Lieman & Israel Pecht

Department of Chemical Immunology, Weizmann Institute of Science, Rehovot 76100, Israel

It has been proposed that protein kinase C mediates cellular responses evoked by external stimuli, leading to alterations in internal free calcium concentrations¹⁻⁴. We have shown previously that histamine-secreting rat basophilic leukaemia cells (RBL-2H3), which degranulate on aggregation of the receptors for immunoglobulin IgE⁵, contain a Ca^{2+} - and phospholipid-dependent protein kinase (kinase C)^{6,7}. The partially purified enzyme is activated directly by the tumour-promoting phorbol ester 12-O-tetradecanoylphorbol-13-acetate (TPA). In intact RBL cells, TPA potentiates histamine release induced⁶ by the Ca^{2+} -ionophore A23187 (similar to the synergy reported for platelets¹, neutrophils² and rat peritoneal mast cells⁸). Although TPA at concentrations below 15 nM synergizes with the antigen, higher TPA concentrations inhibit secretion⁶. This selective inhibition suggested that kinase C is involved in both the activation and termination of the secretory process. To examine this possibility, we have determined the effect of TPA on changes in free cytosolic Ca^{2+} concentration during antigen-induced release. We report here that TPA completely blocks the increase in Ca^{2+} concentration induced by antigen. Our results strongly suggest that protein kinase C is involved in the regulation of receptor-dependent Ca^{2+} signalling.

Using emission of the fluorescent Ca^{2+} indicator quin-2 (refs 6, 9) to monitor changes in the internal concentration of free Ca^{2+} ions in RBL cells, we observed a rapid increase in the internal Ca^{2+} concentration, from an initial value of 93 ± 5 nM, to 350 ± 94 nM (average of 15 different experiments), on stimulation with antigen (Fig. 1a). Maximal increase was reached within

1 min and the Ca^{2+} concentration had begun to decay 3 min later, reaching basal levels ~ 80 min after stimulation. In spite of the rapidity of the increase in internal free Ca^{2+} concentration, it was about 45 min before the histamine secretory process was complete (Fig. 2), indicating that the rate-limiting step in the events leading to degranulation lies distal to Ca^{2+} influx. Addition of TPA alone (1.5-75 nM) failed to trigger any change in the internal Ca^{2+} concentration (Fig. 1b). It did, however, completely block the antigen-induced Ca^{2+} signal (Fig. 1b).

When added after antigen, TPA rapidly reversed the Ca^{2+} signal (Fig. 1c), supporting the notion proposed by Beaven *et al.*¹⁰ that the Ca^{2+} signal is maintained dynamically by the balance between Ca^{2+} influx and active Ca^{2+} efflux. The TPA-induced inhibition of Ca^{2+} influx and the continuous efflux results in reversal of the Ca^{2+} signal. Inhibition of the antigen-dependent Ca^{2+} influx was observed at very low TPA concentrations (1 ng ml^{-1} , 1.5 nM). In contrast, low concentrations of TPA ($<10 \text{ ng ml}^{-1}$, 15 nM), rather than inhibiting antigen-induced release, potentiated it⁶. We therefore attempted to resolve this apparent discrepancy by examining the effect of TPA on antigen-induced release as a function of time after stimulation by antigen. We found that at 7.5 nM, TPA, reported to synergize with antigen-induced histamine release⁶, markedly inhibited release for up to 25 min. This effect was, however, relieved and release was further potentiated 30 min later (Fig. 2). In contrast, the inhibition exerted by a higher concentration of TPA (75 nM) was observed throughout the time range studied (Fig. 2). Significantly, when this high concentration of TPA was added to the cells 20 min after their stimulation with antigen, no inhibition took place (Fig. 2); in fact, release was again potentiated by TPA (Fig. 2). Hence, although in such experimental conditions TPA reverses the Ca^{2+} signal (Fig. 1c), it does not inhibit histamine release.

Protein kinase C has been shown to be the cellular receptor for TPA¹¹. Thus, our results clearly demonstrate that kinase C is involved in an early step of the coupling between stimulus and secretion—the regulation of the plasma membrane permeability to Ca^{2+} ions. When activated directly (by TPA), before exposure to antigen, kinase C inhibits the antigen-induced rise in internal Ca^{2+} concentration. It is possible that this is achieved by effecting closure of the Ca^{2+} gates. Indeed, as these Ca^{2+}

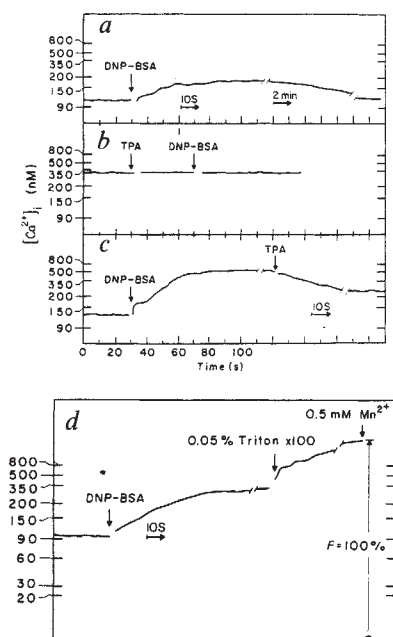


Fig. 1 Fluorescence changes induced by antigen and TPA in RBL-2H3 cells loaded with quin 2, reflecting internal Ca^{2+} concentration, $[\text{Ca}^{2+}]_i$. *a*, Response of the cells to 5 ng ml^{-1} dinitrophenyl bovine serum albumin (DNP-BSA). *b*, Effect of TPA (1 ng ml^{-1} , 1.5 nM) on the internal Ca^{2+} concentration and the antigen-dependent Ca^{2+} signal. *c*, Effect of TPA (50 ng ml^{-1} , 75 nM) on Ca^{2+} signal when added 2 min after antigen. *d*, Effect of lysing the cells with 0.05% Triton X-100 followed by the addition of Mn^{2+} (0.5 mM). *F* represents the percentage of fluorescence obtained. These traces are the results of a typical experiment. Similar results were obtained in 15 similar experiments.

Methods. Cells were maintained in 250-ml tissue culture flasks in Eagle's minimal essential medium (Gibco) supplemented with 15% heat-inactivated fetal calf serum (Biological), 4 mM glutamine and antibiotics in 5% CO_2 at 36°C . Attached cells were collected 48 h after passage by repeated pipetting and suspended in Tyrode buffer containing 137 mM NaCl, 2.7 mM KCl, 0.4 mM NaH_2PO_4 , 2 mM CaCl_2 , 1 mM MgCl_2 , 20 mM HEPES, 5.6 mM glucose and 0.1% BSA pH 7.4. The cells ($1 \times 10^6 \text{ ml}^{-1}$) were incubated with a monoclonal, DNP-specific, IgE antibody for 1 h at 37°C . After several washings with Tyrode buffer, several batches of cells (each 5×10^7 cells ml^{-1}) were subsequently incubated with $100 \mu\text{M}$ of quin 2-acetoxy methyl ester (AME) ($16 \mu\text{l}$ of a 10 mM stock solution in dimethyl sulphoxide to 1.6 ml) for 15 min at 37°C . The cells were also supplemented with 1,5-bis(4-allyldimethylammonium phenyl)pentan-3-one dibromide (BW284C51, Wellcome) to minimize extracellular quin-AME hydrolysis by membranous or secreted esterases. Aliquots of the cell suspension (2×10^6 cells) were then centrifuged and the supernatants discarded. The cell pellets were resuspended to a final volume of 2 ml each and transferred to 1-cm optical path quartz cuvettes, continuously stirred by an inserted electrical stirrer. Fluorescence changes were followed using a thermostated (37°C) Perkin-Elmer MPF-44A spectrofluorimeter (excitation 339 nm, emission 492 nm). Hydrolysis of quin 2-AME to quin 2 and its uptake into the cells were examined by following the shift in emission maximum from 435 to $492 \text{ nm}^{9,10}$.

Received 13 August; accepted 23 October 1984.

1. Sano, K., Takai, Y., Yamanishi, J. & Nishizuka, Y. *J. biol. Chem.* **258**, 2010-2013 (1983).
2. Helfman, D. M., Appelbaum, B. D., Vogler, W. R. & Kuo, J. F. *Biochem. biophys. Res. Commun.* **111**, 847-853 (1983).
3. Knight, D. E. & Baker, P. E. *FEBS Lett.* **160**, 98-100 (1983).
4. Kojima, I., Lippes, H., Kojima, K. & Rasmussen, H. *Biochem. biophys. Res. Commun.* **116**, 555-562 (1983).
5. Segal, D. M., Taurog, J. D. & Metzger, H. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2993-2997 (1977).
6. Sagi-Eisenberg, R. & Pecht, I. *Immun. Lett.* **8**, 237-241 (1984).
7. Sagi-Eisenberg, R., Mazurek, N. & Pecht, I. *Molec. Immun.* (in the press).

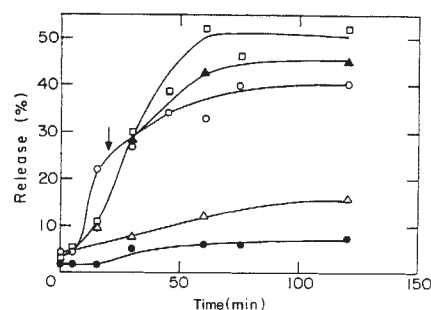


Fig. 2 Effect of TPA on antigen-induced release. ●, Basal release; ○, release induced by 1 ng ml^{-1} DNP-BSA; □, release in the presence of 1 ng ml^{-1} DNP-BSA and 7.5 nM TPA; △, release in the presence of 1 ng ml^{-1} DNP-BSA and 75 nM TPA; ▲, release in the presence of 1 ng ml^{-1} DNP-BSA and 75 nM TPA, with TPA added 20 min after antigen (arrow).

Methods: During the 1 h incubation with IgE, the cells were also supplemented with ^3H -serotonin ($5 \mu\text{Ci ml}^{-1}$, Amersham). After washing, release was induced by adding $100 \mu\text{l}$ cell suspension aliquots to each well of a 96-well flat-bottom microtitre plate (Nunc) containing the various reagents. At the end of the indicated time periods, the cells were spun down and $100 \mu\text{l}$ aliquots of their supernatants counted. The total radiolabelled serotonin incorporated into the cells was determined by lysing the cells in $50 \mu\text{l}$ of 2 M NaOH and counting the radioactivity in the whole sample. Release is expressed as percentage of the total ^3H -serotonin content.

gates are regulated in a potential-independent manner^{12,13}, changes in their state of phosphorylation may be associated with their mode of regulation¹⁴. Alternatively, kinase C may be activating an ATPase responsible for pumping out Ca^{2+} ions. Kinase C is probably also involved in a later step, distal to Ca^{2+} influx; when it is activated, its participation in this step leads to potentiation of the secretory process. Hence, activation of kinase C after stimulation results in enhancement of release. Thus, when TPA is introduced before stimulation, it induces both off and on signals. In the case of treatment with low TPA concentrations ($<15 \text{ nM}$), the potentiating effect of TPA seems to overcome the inhibitory effect, resulting in secretion in the absence of a significant rise in internal Ca^{2+} . Indeed, whereas phorbol esters synergize with internal Ca^{2+} to attain a full cellular response in some cell types^{8,15}, they block response in others^{16,17}. Thus, this apparent dual role of kinase C is not confined only to RBL cells but may represent a universal role for this enzyme in the mediation of cellular responses.

After completion of this manuscript, we were informed that Dr James C. Metcalfe and his colleagues have been conducting similar experiments and eventually obtained the same results.

R.S.-E. is a recipient of the Bat-Sheva De Rothschild foundation grant for the encouragement of Science and Technology, Jerusalem, Israel. This research was supported by a Grant in Leukemia Research by the Leukemia Research Foundation, Inc., Chicago, Illinois, USA.

8. Katakami, Y., Kaibuchi, K., Sawamura, M., Takai, Y. & Nishizuka, Y. *Biochem. biophys. Res. Commun.* **121**, 573-578 (1984).
9. Tsien, R. Y. *Biochemistry* **19**, 2396-2404 (1980).
10. Beaven, M. A. et al. *J. biol. Chem.* **259**, 7129-7136 (1984).
11. Parker, P. J., Stabel, S. & Waterfield, M. D. *EMBO J.* **3**, 953-959 (1984).
12. Sagi-Eisenberg, R. & Pecht, I. *EMBO J.* **3**, 497-500 (1984).
13. Mazurek, N., Schindler, H., Schurholz, T. H. & Pecht, I. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6841-6845 (1984).
14. Sagi-Eisenberg, R. *Trends pharmac. Sci.* (in the press).
15. Michell, B. *Trends biochem. Sci.* **8**, 1-2 (1983).
16. Corvera, S. & Garcia-Sainz, A. *Biochem. biophys. Res. Commun.* **119**, 1128-1133 (1984).
17. Mukhopadhyay, A. K., Bohnet, M. G. & Leidenberger, F. A. *Biochem. biophys. Res. Commun.* **119**, 1062-1067 (1984).

Functional expression of *HLA-DP* genes transfected into mouse fibroblasts

Penelope Austin*, John Trowsdale*, Christopher Rudd*, Walter Bodmer*, Marc Feldmann† & Jonathan Lamb†

* Imperial Cancer Research Fund, Lincoln's Inn Fields, PO Box 123, London WC2A 3PX, UK

† Imperial Cancer Research Fund, Tumour Immunology Unit, Department of Zoology, University College London, Gower Street, London WC1E 6BT, UK

The HLA class II antigens are a highly polymorphic family of dimeric cell-surface glycoproteins, expressed predominantly on the surface of immunocompetent cells. They are intimately involved with the induction of the T-cell response to extrinsic antigen¹⁻⁴ and are important predisposing factors for a wide spectrum of autoimmune diseases⁵. We describe here the expression of a class II product from the *HLA-DP* (new WHO nomenclature, formerly *SB*) subregion after transfer of cloned genes into mouse fibroblasts. The transfected DP antigen is recognized by several HLA class II monoclonal antibodies and, though present in a mouse cell background, is able to function in the presentation of influenza antigen to cloned DP-restricted human T lymphocytes.

Molecular cloning studies show that the HLA class II region contains a minimum of six α -chain⁶ and seven β -chain genes, which on sequence analysis can be resolved into the three subgroups—*DP* (formerly *SB*), *DQ* (formerly *DC*) and *DR*—defined by serological and cellular typing (for a review, see ref. 4). The *DP* region, originally defined by primed lymphocyte typing⁷, contains two α and two β genes arranged in the order *DP β 2*, *DP α 2*, *DP β 1*, *DP α 1* (ref. 8). The products of this complex, closely linked and highly related gene family are normally expressed coordinately, posing problems in attributing structural features, such as antigenic epitopes, and functional properties, such as T-cell recognition, to any one antigen. The transfection and expression of cloned genes is thus a useful approach to characterizing the structure and function of each product in isolation.

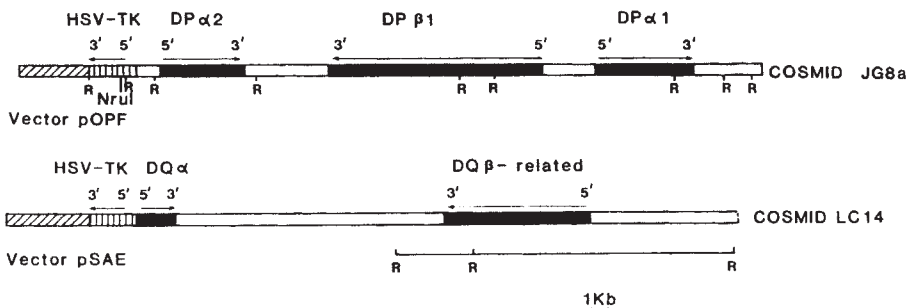
Expression after gene transfer of both human and mouse major histocompatibility complex (MHC) class II genes has been reported⁹⁻¹². In the mouse, the functional activity of the transfected molecules was also investigated and the capacity of B-cell lymphomas¹⁰, interferon-induced macrophages¹¹ and L-cell fibroblasts¹² to present antigen after transfection was demonstrated. The B-cell transfectants were also shown to effect allostimulation¹⁰ and the L-cell transfectants to act as targets for allogeneic killing¹². The present experiments aimed to demonstrate the expression of human *HLA-DP* genes in mouse L cells and to determine their ability to activate DP-restricted influenza virus-specific T cells.

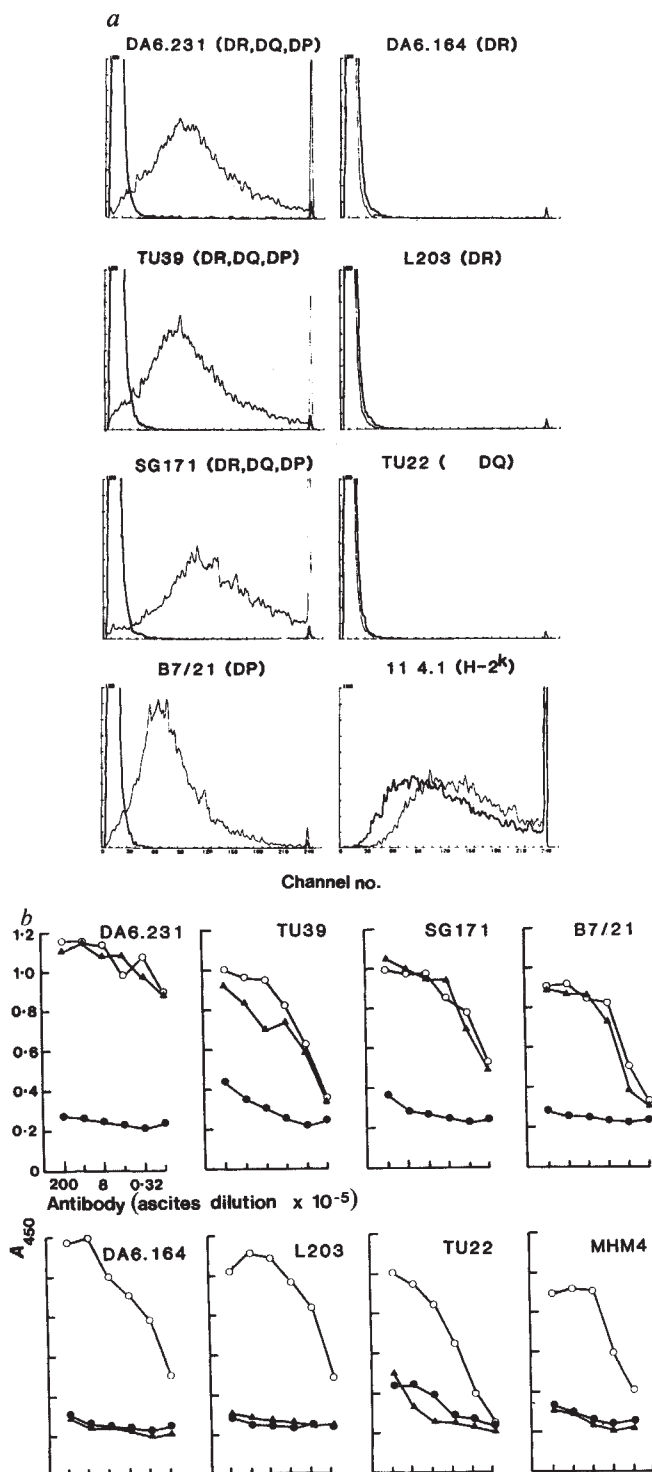
The cosmid clone JG8a, isolated from a library of placental DNA, contains two *DP α* -related genes and one *DP β* -related gene⁸ (Fig. 1). The genes were introduced into mouse L cells by co-precipitation with calcium phosphate followed by selection for thymidine kinase (TK). An initial analysis of TK⁺ transfectants for expression at the messenger RNA level showed that transcription was occurring from all three genes (data not shown). On testing mass cultures of JG8a transfectants with a variety of HLA class II monoclonal antibodies, DA6.231 (anti-DR, -DQ and -DP) gave consistent, above-background staining on flow microfluorimetry analyses using the fluorescence-activated cell sorter (FACS). This antibody was then used to select for transfectants with comparatively high levels of DP expression, either by FACS sorting or screening of primary TK⁺ transfectants. The experiments described here analyse the antibody-binding patterns and functional activity of one such selected line.

The pattern of antibody binding to the DP transfectants is illustrated by the data from FACS analysis and from an enzyme-linked immunosorbent assay (ELISA) (Fig. 2a, b). The results show positive binding of four HLA class II monoclonal antibodies, DA6.231 (anti-DR, -DQ and -DP), TU39 (anti-DR, -DQ and -DP), SG171 (anti-DR, -DQ and -DP) and B7/21 (anti-DP, originally anti-FA, see below and ref. 13), but only background binding with DA6.164 (anti-DR), L203 (anti-DR), TU22 (anti-DQ, possibly also anti-DR) and MHM4 (classified as anti-DP). The locus specificity attributed to these antibodies is based on binding studies of γ -irradiation-induced B-cell deletion mutants selected for loss of class II expression^{14,35} and immunochemical analysis on two-dimensional gels^{15,16}. In the case of DA6.231,

Fig. 1 Schematic maps and transfection of cosmids JG8a and LC14. Cosmid JG8a contains two DP-region α -chain genes and one β -chain gene. The restriction sites for *Eco*RI and *Nru*I are shown. Further details have been published elsewhere⁸. Cosmid LC14²⁶ contains an incomplete *DQ α* gene and a *DQ β* -related gene found on *Eco*RI fragments of 5.5 and 18 kb as shown.

Methods: DNA was introduced into a subline of LTK⁻ cells previously shown to give good surface expression of transfected *HLA-DR* genes⁹. The technique adopted was that of co-precipitation with calcium phosphate²⁷ using purified cosmid DNA and LTK⁻ carrier DNA at a final concentration of 15 μ g ml⁻¹ transfection cocktail. The precipitate was added to LTK⁻ cells, seeded at a density of 0.5 $\times$ 10⁶ cells per 9-cm dish, and given a fresh change of Dulbecco's modified Eagle's medium (DMEM) with 10% fetal calf serum (FCS) 4 h before transfection. 0.66 ml of transfection cocktail was added to 5.5 ml medium in each dish and after 4–10 h incubation at 37 °C, 10% CO₂, a glycerol shock was given. The cells were overlaid with 25% glycerol in DMEM for 1 min before washing and adding 10 ml DMEM with 10% FCS and 5 mM sodium butyrate²⁸. After a further 12 h incubation, the medium was changed to selective medium containing 15 μ g ml⁻¹ hypoxanthine, 0.2 μ g ml⁻¹ methotrexate and 5 μ g ml⁻¹ thymidine. The cells were subsequently maintained in this medium with changes every 3–4 days. TK⁺ colonies were visible after ~12 days and were either removed by trypsinization to form a mass culture or trypsinized individually using cloning rings. DNA used in the transfection experiments to produce the lines described was as follows. L8a.5: 5 μ g JG8a cut with *Nru*I, 1 μ g pOPF cut with *Eco*RI, 5 μ g LTK⁻ DNA; LpOPFm: 1 μ g pOPF cut with *Eco*RI, 9 μ g LTK⁻ DNA; L14m: 2 μ g LC14, 8 μ g LTK⁻ DNA. The co-transfection procedure used to produce L8a.5 reduced the efficiency of TK⁺ colony formation at least 10-fold whilst improving the levels of DP expression observed in the resulting transfectants. The *Nru*I digestion linearizes the cosmid and destroys thymidine kinase activity, while cutting at the *Eco*RI site at the 5' end of the *tk* gene truncates important regulatory sequences and reduces transcription levels. Use of this *Eco*RI fragment has been described in constructs to promote high copy number integration²⁹. The line L8a.5 was derived from a single TK⁺ colony and selected for further study on the basis of comparatively high levels of DP expression. Southern blot analyses demonstrated the high copy number of integrated sequences found (data not shown). The lines LpOPFm and L14m are mass cultures derived from ~100 colonies and are used as controls in subsequent experiments. Before experimental analysis, the L cells were incubated for 16–24 h in medium containing 5 mM sodium butyrate, a procedure found to enhance levels of DP expression.





TU39, SG171, TU22, L203 and DA6.164, the binding results for the DP transfectants strengthen the argument for these locus assignments. Positive binding of B7/21 supports the view that this antibody detects DP-region molecules¹⁷, previously termed FA¹². Conversely, the lack of binding by the MHM4 antibody questions its specificity for DP, although it may well recognize the DP β 2 product or a polymorphic DP β 1 product, not expressed by the transfectants. These results emphasize the value of such transfectant lines in elucidating the complex cross-reactions of HLA class II antibodies. Accurate quantitation of the level of surface antigen expression relative to that seen on B-cell lines is difficult because of these cross-reactions. However, comparison of mean fluorescence levels observed on binding B7/21 to the B-cell line LNAT suggests that surface expression is approximately fivefold lower on the DP transfectants.

Fig. 2 Surface expression of DP antigen by transfected mouse L cells. *a*, Flow microfluorimetry analyses of the transfectants L8a.5 (thin lines) and L14m (thick lines) using a panel of monoclonal antibodies to HLA class II determinants and an anti-H-2^k class I antibody. Fluorescence intensity is plotted as a linear function of cell number and 10,000 cells were analysed for each histogram. *b*, ELISA assay for binding of HLA class II monoclonal antibodies to the EBV-transformed lymphoblastoid cell line LNAT (○) and L-cell transfectants Lp0PFm (●) and L8a-5 (▲).

Methods: *a*, 1×10^6 cells were washed with cold phosphate-buffered saline (PBS) containing 5% fetal calf serum (FCS) and incubated with 100 μ l of antibody diluted in PBS+5% FCS for 30 min on ice. All HLA class II antibodies were used in the form of ascites fluid diluted 1 in 200; 11.4.1 was used as a hybridoma supernatant diluted 1 in 2. Cells were washed three times in cold PBS+5% FCS and incubated for a further 30 min on ice with a 1 in 10 dilution of fluoresceinated IgG fraction rabbit anti-mouse IgG (Cappel Laboratories). After two final washes the cells were resuspended in cold PBS and analysed on a FACS-1 (Becton & Dickinson). *b*, Cells were washed three times in PBS and added to 96-well plates pretreated for 1 h with 0.1 mg ml⁻¹ poly-L-lysine in PBS (10^5 cells in 50 μ l PBS per well). The cells were spun onto the base of the wells and then fixed by gently adding 200 μ l per well 0.025% glutaraldehyde in PBS and standing at room temperature for 20 min. After washing, the plates were stored at 4 °C in a solution of 200 μ g ml⁻¹ gelatin, 0.02% NaN₃ in PBS. Before adding antibody, the cells were treated with 0.1% phenylhydrazine in PBS for 1 h and then washed in PBS (3 $\times$) and once in PBS+0.2% Tween 20. Antibody ascites, diluted at 1 in 500 in PBS+5% FCS and at subsequent five-fold serial dilutions, were added to the cells at 50 μ l per well in triplicate and the plates were incubated at room temperature for 2 h. The cells were washed (as above) and then incubated for 30 min with 25 μ l per well 5% goat anti-mouse immunoglobulin (Central Laboratory of the Netherlands Red Cross Blood Transfusion Service) in 200 mM Tris-HCl pH 7.6 with 10% normal goat serum (NGS). After further washes the plates were incubated for 30 min with peroxidase-antiperoxidase immune complexes (PAP) in Tris-HCl 7.6+10% NGS. After final washing, substrate (1 mg ml⁻¹ *o*-phenylenediamine, 0.05% H₂O₂ in substrate buffer, 156 mM Na₂HPO₄, 27 mM citric acid pH 6.0) was added (100 μ l per well) and the plates were incubated in darkness for 30 min. Absorbance at 450 nm was read on the Titer Multiscan.

Results are expressed as the mean of triplicate wells.

The presence of the DP antigen was confirmed biochemically by immunoprecipitation using the monoclonal antibodies SG171 and B7/21 coupled with two-dimensional non-equilibrium pH gradient electrophoresis/sodium dodecyl sulphate polyacrylamide gel electrophoresis (NEPHGE/SDS-PAGE) analysis. Figure 3 shows the two-dimensional pattern of ¹²⁵I-labelled polypeptides derived from microsomes of the mouse L cells (Fig. 3a) and the L cells transfected with cosmid clone JG8a (Fig. 3b). In addition to the nonspecific polypeptides found in both cells, the monoclonal antibodies precipitated a single putative α -chain of relative molecular mass (M_r) 30,000–32,000 and at least two possible β subunits of 29,000–32,000 from the DP-transfected cells. The identification of these spots is based on the isoelectric points and M_r positions previously found for HLA-D region antigens^{17,18}. The labelled α -chain spot suggests

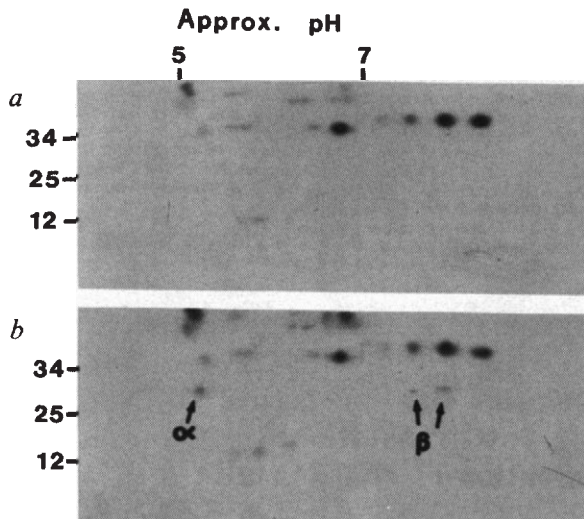


Fig. 3 Two-dimensional NEPHGE/SDS-PAGE analysis of immunoprecipitates using a combination of monoclonal antibodies SG171 and B7/21 against mouse L cells (*a*) and mouse L cells transfected with cosmid clone JG8a (*b*). Numbers on the left-hand side represent size markers ($M_r \times 10^{-3}$).

Methods: Microsomes were prepared by shearing the cells in a Stansted cell disrupter, centrifuging at 1,500 r.p.m. for 5 min to remove nuclear debris and finally at 18,000 r.p.m. for 30 min to pellet the membrane vesicles. The resuspended pellet was labelled with ^{125}I by lactoperoxidase-catalysed iodination³², washed and then solubilized in lysis buffer (1% (w/v) Nonidet P-40 in 10 mM Tris-HCl buffer pH 8.2 containing 150 mM NaCl, 1 mM EDTA and 1 mM phenylmethylsulphonyl fluoride). A glycoprotein fraction was prepared by binding to *Lens culinaris* lectin and eluting with 10% α -methyl-D-mannoside. The lysate (5×10^7 cell equivalents) was incubated overnight at 4°C with 5 μl of ascitic fluid of SG171 and B7/21 before precipitating the immune complexes with 20 μl of 10% *Staphylococcus aureus* Cowan 1 strain (SaC1). Immunoprecipitates were washed twice in lysis buffer containing 500 mM NaCl and once in lysis buffer alone. Immune complexes were eluted from SaC1 pellets by incubating in isoelectric focusing sample buffer (9.2 M urea, 2% Nonidet P-40, 5% ampholines pH range 3.5–10 (LKB) at 50°C for 30 min. Two-dimensional NEPHGE/SDS-PAGE analysis was done as described previously^{30,31}.

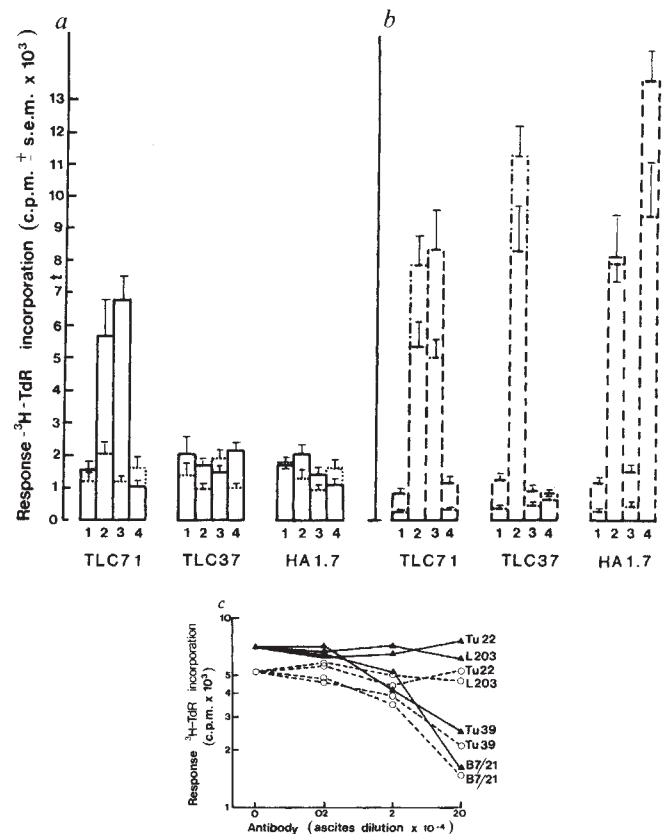


Fig. 4 Antigen presentation assay. The proliferation response of T-lymphocyte clones TLC71 (neuraminidase specific, DP-restricted), TLC37 (matrix protein specific, DR-restricted) and HA1.7 (peptide 20 specific, DQ-restricted) in the absence of antigen (column 1), or with intact virus (A/Texas/1/77 $\times 49$ at 5 haemagglutinating units ml^{-1} , column 2), or neuraminidase (N_2 of Papua New Guinea/1/75 at 10^{-3} v/v, column 3), or p20 ($1.0 \mu\text{g ml}^{-1}$, column 4). Presenting cells: *a*, DP L-cell transfectants L8a.5 (—) ($1.25 \times 10^5 \text{ ml}^{-1}$); DQ L-cell transfectants L14m (· · · · ·) ($1.25 \times 10^5 \text{ ml}^{-1}$). *b*, Autologous PBL (— · — · —) ($1.25 \times 10^5 \text{ ml}^{-1}$); EBV-transformed autologous B-cell line LNAT (— · — · —) ($0.5 \times 10^5 \text{ ml}^{-1}$). *c*, Inhibition of the proliferation response by TLC71 with anti-HLA Class II antibodies. The assay for the proliferative response to intact virus was performed as before using DP transfectants L8a.5 (Δ) and autologous PBL ($\circ$) as presenting cells. Antibody ascites was added at the initiation of the cultures in the dilutions shown, and left in for the duration of the experiment.

Methods: The isolation and characterization of T-lymphocyte clones TLC71, HA1.7 and TLC37 have been described in detail elsewhere^{19,20,33,34}. Briefly, PBL from the donor LNAT were cultured for 6 days with specific antigen. The lymphoblasts were enriched on a discontinuous Percoll (Pharmacia) gradient and resuspended in RPMI-1640 (Gibco) containing 10% A^+ serum and 20% interleukin-2 (IL-2) and plated at one cell every third well in Microtest II trays (Falcon) with 10^4 irradiated (2,500 rad) autologous PBL and antigen. After 7 days, growing clones were transferred to 96-well microtitre trays and subsequently to 24-well trays. At each transfer the clones received fresh IL-2 and irradiated PBL together with antigen. The clones were expanded in 25- cm^2 tissue culture flasks receiving IL-2 every 3 days and irradiated histocompatible PBL and intact virus (A/Texas/1/77) every 7 days. IL-2 was prepared from 48-h supernatants of PBL ($1 \times 10^6 \text{ ml}^{-1}$) cultured with 0.1% purified phytohaemagglutinin (PHA-P, Difco) in complete medium containing 2.5% A^+ serum. The clones were used in proliferation assays 6–7 days after the addition of feeder cells (irradiated PBL). Cloned T cells ($5 \times 10^4 \text{ ml}^{-1}$) were cultured together with the antigens and presenting cells described above. The L-cell transfectants and EBV-line received 5,000 rad and PBL 3,000 rad before use. Following 60 h incubation, the cultures were pulsed for 8–16 h with $1.0 \mu\text{Ci } ^3\text{H}$ thymidine (^3H -TdR, Amersham) and collected onto glass fibre filters. Proliferation as correlated with ^3H -TdR incorporation was measured by liquid scintillation spectroscopy. Results are expressed as the mean counts per min of triplicate cultures $\pm$ s.e.m.

that, contrary to the case of the DP α -chain on the surface of intact B cells^{13,17}, the α -chain in microsomes from the L-cell transfectants is accessible to lactoperoxidase-catalysed iodination. It is likely to correspond to the DP α 1 gene product, as the sequence from this gene matches the sequence available for DP α cDNA clones and protein sequence data⁸. However, the data do not exclude concomitant expression of the DP α 2 gene. The presence of multiple β -chains probably reflects stages of glycosylation¹⁸.

Having established that the JG8a transfectants were expressing a DP antigen at the cell surface, the L-cell transfectants were assayed for their ability to induce proliferation of the T-lymphocyte clone TLC71 in the presence of antigen. TLC71 has been shown to proliferate in response to DPw2 class II determinants¹⁹ and the neuraminidase (NA) glycoprotein of influenza A virus²⁰. In the presence of intact influenza virus (A/Texas/1/77) or NA (N_2 of Papua New Guinea/1/75), the DP L-cell transfectant was able to induce TLC71 to proliferate (Fig. 4a). However, in the presence of an irrelevant antigen, peptide 20 (p20; residues 306–329 of the HA1 molecule of influenza haemagglutinin²¹) or with DQ transfectants, there was no response. Neither HA1.7 (DQ restricted and specific for p20) nor TLC 37 (DR restricted and specific for matrix protein) T-cell clones were activated by the L-cell transfectants. However, both HA1.7 and TLC 37 clones proliferated in response to specific antigen presented by autologous peripheral blood lymphocytes (PBL) or Epstein-Barr virus (EBV)-transformed B

cells (Fig. 4b). These results show that T-cell activation by the DP transfectants is both *HLA* dependent and antigen specific.

To confirm the specificity of the transfected DP molecule in antigen-induced activation of TLC71, the inhibitory effect of antibodies to various class II *HLA* molecules was tested. Figure 4c shows that the anti-DP antibodies TU39 and B7/21 inhibit the proliferation of TLC71 whereas TU22 and L203 have no effect. In a parallel experiment, TU22 (anti-DQ) and L203 (anti-DR) effectively inhibited the response of HA1.7 (DQ-restricted) and TLC37 (DR-restricted), respectively, to intact A/Texas in the presence of autologous PBL (data not shown). Further analysis of the polymorphic specificity of T-cell recognition can now be pursued using both transfectants and T-cell clones of different specificities.

The capacity of the DP-expressing mouse fibroblasts to activate human T cells in the presence of antigen indicates that antigen presentation is not restricted to the macrophage/monocyte or B-cell lineages. This agrees with our recent observation that thymocytes expressing class II antigens can activate human T-cell clones²² and also with the results for the mouse class II L-cell transfectants¹². The implication is that either secondary signals such as the secretion of interleukin-1 (IL-1) are not always essential or that cells other than classical antigen-presenting cells such as mouse fibroblasts are able to provide them. Similarly, the postulated interactions between the T-cell surface antigens such as T4 (refs 23, 24) and LFA-1 (ref. 25) and the antigen-presenting cell are either not obligatory or are directed against targets present on the surface of the L-cell transfectants. Requirements for T-cell activation may well be more stringent for primary rather than secondary stimulation and may also vary with the particular antigen and T-cell clone involved. However, in the system described, it is clear that antigen-dependent activation can occur in the absence of any human cell-surface molecule on the presenting cell, other than the *HLA* restriction element recognized by the T-cell receptor.

We thank Dr C. du Preval for the gift of Ltk⁻ cells, Drs K. Guy and V. Van Heyningen for DA6.231 and DA6.164, Dr I. Trowbridge for B7/21, Dr A. Ziegler for TU39 and TU22, Dr S. Goyert for SG171, Dr A. McMichael for MHM4, Dr L. Lampson for L203, Dr A. Rosenthal (Merck Sharp and Dohme) for A/Texas/1/77, Dr R. Webster for influenza virus neuraminidase, Dr R. Lerner for peptide 20, Dr M. Contreras and her associates at the National Blood Transfusion Centre, Edgware for blood and serum, Grace Lam and Laura Davis for their assistance in operation of the FACS, Margaret Hollis and Audrey Symons for preparation of the manuscript, and Professor N. A. Mitchison and Professor Martin Raff for helpful discussion and comments.

Received 30 July; accepted 3 October 1984.

- Klein, J. & Nagy, Z. A. *Adv. Cancer Res.* **37**, 233-317 (1982).
- Bodmer, W. F. in *Mammalian Genetics and Cancer* (ed. Russel, E.) 213-240 (Liss, New York, 1981).
- Schwartz, R. H. in *In Antigens* (eds Ferrone, S. & David, C. S.) 161-218 (CRC Press, Boca Raton, 1982).
- Kaufman, J. F., Auffray, C., Korman, A. J., Shackelford, D. A. & Strominger, J. *Cell* **36**, 1-13 (1984).
- Svejgaard, A., Platz, P. & Ryder, L. P. *Immun. Rev.* **70**, 193-218 (1983).
- Spielman, R., Lee, J. S., Bodmer, W. F., Bodmer, J. G. & Trowsdale, J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3461-3465 (1984).
- Shaw, S., Pollack, M. S., Payne, S. M. & Johnson, A. H. *Hum. Immun.* **1**, 177-185 (1980).
- Trowsdale, J. *et al. Cell* **38**, 241-249 (1984).
- Rabourdin-Combe, C. & Mach, B. *Nature* **303**, 670-674 (1983).
- Germain, R. N., Norcross, M. A. & Margulies, D. H. *Nature* **306**, 190-194 (1983).
- Foisom, V. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 2045-2049 (1984).
- Malissen, B. *et al. Cell* **36**, 319-327 (1984).
- Watson, A. J., DeMars, R., Trowbridge, I. S. & Bach, F. C. *Nature* **304**, 358-361 (1983).
- Shaw, S., Ziegler, A. & DeMars, R. *Hum. Immun.* (in the press).
- Bodmer, J. G. & Bodmer, W. F. *Br. med. Bull.* (in the press).
- Crumpton, M. J. *et al. in Histocompatibility Testing 1984* (eds Albert, E. & Mayr, W.) (Springer, New York, in the press).
- Rudd, C. E., Bodmer, J. G., Bodmer, W. F. & Crumpton, M. J. in *Histocompatibility Testing 1984* (eds Albert, E. & Mayr, W.) (Springer, New York, in the press).
- de Kretser, T. A., Crumpton, M. J., Bodmer, J. G. & Bodmer, W. F. *Molec. biol. Med.* **1**, 59-76 (1983).
- Eckels, D. D. *et al. Nature* **301**, 716-718 (1983).
- Lamb, J. R., Eckels, D. D., Phelan, M. & Woody, J. N. *J. Immun.* **128**, 1428-1432 (1982).
- Green, N. *et al. Cell* **28**, 477-488 (1982).
- Londei, M., Lamb, J. R., Bottazzo, G. F. & Feldmann, M. *Nature* (in the press).
- Meuer, S. C. *et al. Science* **218**, 471-473 (1982).

- Biddison, W., Rao, P. E., Talle, M. A., Goldstein, G. & Shaw, S. *J. Immun.* **131**, 152-157 (1983).
- Springer, T. A. *et al. Immun. Rev.* **63**, 171-195 (1982).
- Trowsdale, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 1972-1976 (1983).
- Wigler, M., Pellicor, A., Silverstein, S. & Axel, R. *Cell* **14**, 725-731 (1978).
- Gorman, C. M., Howard, B. H. & Reeves, R. *Nucleic Acids Res.* **11**, 7631-7649 (1983).
- Gillies, S. D., Morrison, S. L., Oi, V. T. & Tonegawa, S. *Cell* **33**, 717-718 (1983).
- O'Farrell, P. L., Goodman, H. J. M. & O'Farrell, P. H. *Cell* **12**, 1133-1142 (1977).
- Rudd, C. E., Bodmer, J. G., Bodmer, W. F. & Crumpton, M. J. *J. biol. Chem.* (in the press).
- Hubbard, A. L. & Cohn, Z. A. *Biochemical Analysis of Membranes* (ed. Maddy, A. H.) 427-501 (Chapman & Hall, London, 1976).
- Lamb, J. R. & Feldmann, M. *Nature* **308**, 72-74 (1984).
- Lamb, J. R., Eckels, D. D., Lake, P., Woody, J. N. & Green, N. *Nature* **300**, 66-69 (1982).
- Shaw, S. & De Mars, R. *Disease Markers* **2**, 183-195 (1984).

Polypeptide ligation occurs during post-translational modification of concanavalin A

D. M. Carrington*, A. Auffret† & D. E. Hanke*

* Botany School, University of Cambridge, Downing Street, Cambridge CB2 3EA, UK

† SERC Protein Sequencing Unit, Department of Biochemistry, University of Leeds, 9 Hyde Terrace, Leeds LS2 9LS, UK

Lectins are proteins with multivalent carbohydrate-binding sites, which confer the ability to agglutinate. The seeds of legumes are particularly rich in lectins, for example, concanavalin A (Con A) comprises up to 15% of the protein in the cotyledons of jack bean (*Canavalia ensiformis*) seeds. The amino acid sequences of Con A and several other legume lectins have been partially or fully determined, and comparison of these sequences from different species reveals a circular homology^{1,2} (Fig. 1A); rearrangements within the genome have been suggested to explain this^{1,2}. We report here that the circular homology displayed by Con A is due to a post-translational transposition and ligation within the initial polypeptide. This type of modification has not been reported previously for eukaryotes, although it has been suggested to occur in bacteriophage λ (ref. 3).

A proportion of the Con A monomer extracted from mature seeds is cleaved after Asn 118 (ref. 4), yielding three bands on SDS-polyacrylamide gel electrophoresis (PAGE) (Fig. 1B). An initial observation that some of the monomer extracted from immature seeds had a slightly higher relative molecular mass (M_r) (Fig. 1B) led us to determine the N-terminal sequences of the two forms (Fig. 1C). The additional form found in immature seeds has an N-terminal extension of four residues, the remainder of the sequence being the same as that reported previously⁵.

A form of Con A having no lectin activity (37CA) was detected using Western blotting (Fig. 2A); this form is found throughout the growth of the seed but disappears on maturation. The pattern of fragments of Con A in immature seeds (Figs. 1B, 2A) is different from that of the mature seed; this difference has not yet been explained. Pulse-chase experiments *in vivo* with radio-labelled substrates showed that 37CA is the precursor of the other forms of Con A (Fig. 2B).

RNA extracted from all stages of developing cotyledons directed protein synthesis in a reticulocyte lysate⁶. The peptides produced by cyanogen bromide (CNBr) cleavage of 37CA and of the immunoprecipitated *in vitro* form of Con A are similar, but both sets of peptides differ from those produced by cleavage of Con A (Fig. 3). This difference was explained only when DNA complementary (cDNA) to the messenger RNA for Con A was sequenced (Fig. 4A): the positions of the methionines are consistent with the CNBr cleavage pattern of 37CA (Fig. 4B). The most striking observation is that the amino acid sequence derived from the cDNA has direct, not circular, homology with other legume lectins, thus in forming mature Con A there must be a transposition and ligation of two peptides

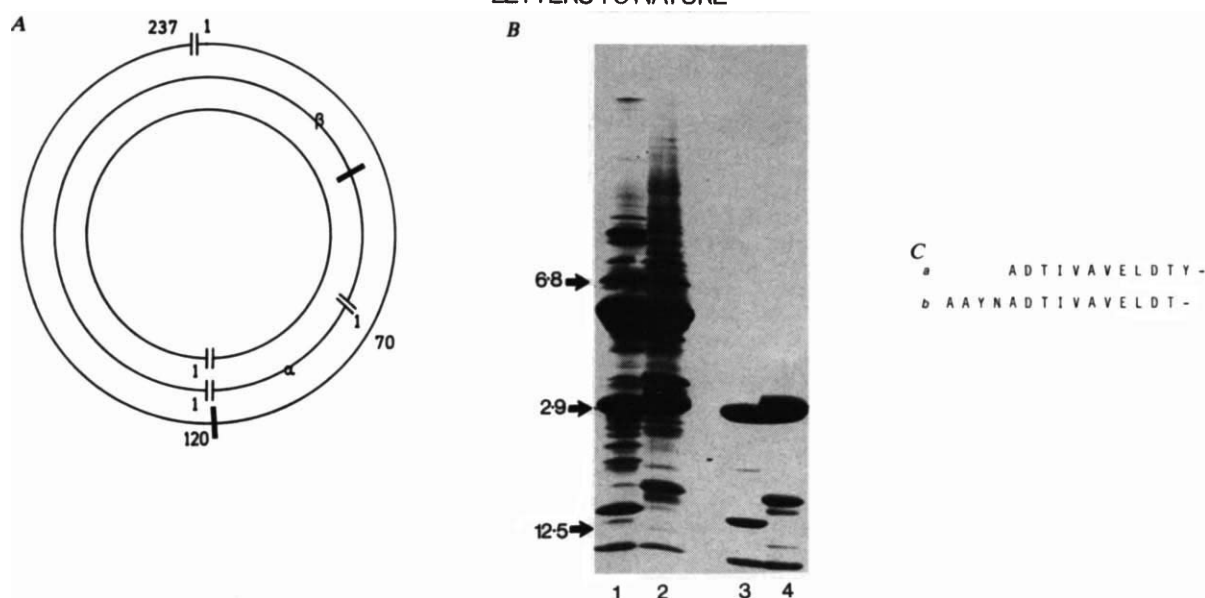


Fig. 1 A, Representation of circular homology in amino acid sequences of various legume lectins^{1,2}. Outer circle, Con A (the bar represents the position of cleavage found in a proportion of the molecules of Con A). Middle circle, favin or lentil lectin (β -chain extends to the bar). Inner circle, soya bean agglutinin and phytohaemagglutinin. B, SDS/7–20% polyacrylamide gel¹¹ of: lane 1, total polypeptides from mature cotyledons; lane 2, total polypeptides from the cotyledons of an immature seed (2 g); lanes 3, 4, forms of Con A purified by affinity chromatography from mature cotyledons and cotyledons of 2-g immature seeds, respectively. To prepare Con A, cotyledons were homogenized in 25 mM MOPS, 125 mM NaCl solution pH 6.8, and the lectin purified by affinity chromatography on Sephadex G-50 (fine); elution was with 100 mM α -methyl mannoside in the homogenization buffer. In the growth conditions used, the seeds grew to a final fresh weight of ± 4 g at ~ 60 days post-pollination before loss of weight by dehydration began. Almost all the increase in fresh weight occurred between 30 and 60 days post-pollination. C, The N-terminal amino acid sequences of Con A from mature (a) and immature (b) seeds. Sequence analysis was performed with a Rank Hilger APS 240 solid-phase sequencer or a Beckman model 890C liquid-phase (spinning cup) sequencer using standard programmes. For solid-phase sequencing the protein was coupled to the *N*-(2-aminoethyl)-3-aminopropyl derivative (AEAP) of 7.5 nm controlled-pore glass¹² by the carbodiimide/hydroxybenzotriazole method¹³. Liquid-phase sequencing used programme 030176 (a Beckman version of the procedure of Brauer *et al.*) and 0.25 M Quadrol buffer in the presence of polybrene¹⁵. Phenylthiohydantoin (PTH) amino acids were identified by reverse-phase HPLC on Zorbax C₁₈ columns.

Fig. 2 A, Polypeptides immunologically related to Con A in the developing cotyledons, visualized after SDS/7–20% PAGE. Lanes 1–7, the forms found in 0.35, 0.6, 1.25, 2.1, 3.4, 4.3-g (all fresh weight) and mature seeds, respectively. The arrows indicate 37CA and the two forms of Con A having lectin activity (Fig. 1B). B, Fluorograph of a SDS-polyacrylamide gel showing a pulse-chase experiment involving the radiolabelling of total cotyledon polypeptides from 1-g seeds incubated with ³H-leucine for 2 h. Lane 1, no chase; lanes 2–4, chases of 2, 8 and 20 h, respectively; lane 5, polypeptides stained with Coomassie brilliant blue. The positions of *M_r* standards are shown to the right. The arrows on the left show the positions of forms and fragments of Con A (see lanes 2 and 3 of A).

Methods: A, Cotyledons were homogenized in 100 mM Tris, 10% glycerol, 2% SDS, 2% mercaptoethanol, 0.1% bromophenol blue solution pH 6.8 (sample buffer) and incubated at 100 °C for 5 min. After electrophoresis the gel was washed with 0.2 M sodium phosphate, 0.1% SDS solution (pH 7.5) for 3 h with several changes before electrophoretic transfer to diazobenzoxymethyl paper¹⁶. The probe used was IgG purified from an antiserum raised against electrophoretically pure Con A (from mature seeds) in a rabbit, followed by affinity-purified ¹²⁵I-mouse anti-rabbit IgG (200 μ Ci μ g⁻¹). B, The testa were removed from the seeds and each pair of cotyledons placed in 1 ml of 2 mM glutamine in Murashige and Skoog's²⁷ nutrient solution containing 200 μ Ci of [4, 5-³H]leucine (120 Ci mmol⁻¹) for 3 h. For the chase, the cotyledons were transferred to 1 ml of the same solution, except that the radiolabel was replaced with 2 mM leucine. Cotyledons were homogenized in 75 mM Tris-HCl, 2% SDS, 10% sucrose, 0.1% bromophenol blue solution pH 6.8, and incubated at 100 °C for 5 min.

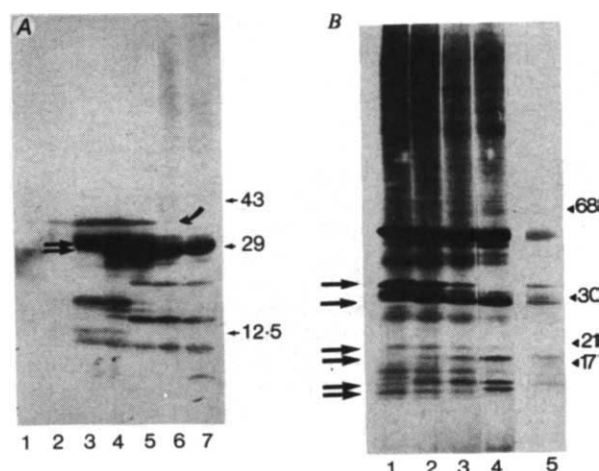
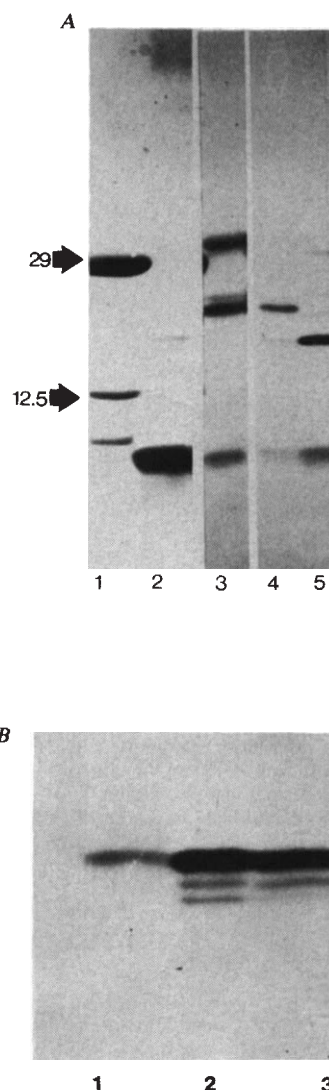


Fig. 3 A, Peptides produced by CNBr cleavage of various forms of Con. A, separated in a 7–20% SDS-polyacrylamide gel. Lane 1, intact Con A; lane 2, Con A extracted from mature cotyledons; lane 3, 37CA; lane 4, nascent Con A synthesized *in vitro* in the presence of canine pancreatic microsomes; lane 5, nascent Con A synthesized *in vitro*. Lanes 1 and 2 were stained with Coomassie brilliant blue; lanes 3, 4 and 5 were fluorographed. The single major band of lane 2 corresponds to two of the expected CNBr fragments (residues 42–128 and 129–237) which migrate closely together. The third fragment (residues 1–41) is too small to appear on this gel (Fig. 4C). The 37CA purified by electroelution does not cleave fully (three separate attempts), which is why the intact monomer appears in lane 3 in addition to fragments of M_r 20,000 (20K) and 9K. The forms of Con A synthesized *in vitro* are immunoprecipitated, not electroeluted, and undergo virtually complete cleavage (lanes 4, 5). Provided canine pancreatic microsomes are present during *in vitro* translation, the immunoprecipitated polypeptides cleave to produce fragments which co-migrate with the 20K and 9K fragments produced from 37CA, consistent with other results (to be published elsewhere), establishing that 37CA is glycosylated. B, Fluorograph of the forms of Con A purified. Lane 1, electroeluted 37CA radiolabelled *in vivo*; lanes 2, 3, the forms immunoprecipitated from *in vitro* translation products using total cotyledon RNA (lane 2) or total cotyledon RNA and canine pancreatic microsomes (lane 3) to direct synthesis. No other polypeptides were visible on the fluorograph. We consider that the minor components immunoprecipitated from the *in vitro* translation products are artefacts of the translation system because: (1) they are reduced in the presence of canine pancreatic microsomes; (2) a different, single, minor component is observed when polysomes are used to direct synthesis; (3) only one band is observed on Northern blots of total cotyledon RNA probed with a cDNA clone of Con A mRNA; (4) Southern blots of genomic DNA indicate that there is only one gene for Con A (to be published).

Methods: A, 10 mM CNBr in 70% formic acid was added to the dry polypeptides to give a ± 500 -fold excess, and the mixture left at room temperature for 24 h, then water (10 vol.) was added and the peptides recovered by lyophilization. ^3H -37CA was prepared from cotyledons radiolabelled *in vivo* with ^3H -leucine¹⁷ by two rounds of electroelution from preparative SDS-polyacrylamide gels¹⁸. Total cotyledon RNA from immature seeds (1.5 g) was used to direct polypeptide synthesis *in vitro*. Total RNA was prepared by a modification of the method of H. R. B. Pelham (unpublished). Cotyledons were removed from seeds and ground to a fine powder under liquid nitrogen. The nitrogen was allowed to evaporate and the powder added to 20 vol. of 50 mM Tris, 150 mM NaCl, 10 mM EDTA, 1% SDS solution pH 7.8 (NETS)/phenol (saturated with NETS) 1:1, at 80 °C. After vigorous shaking, the mixture was kept at 80 °C for 5 min, then cooled to room temperature. The phases were separated after centrifugation and the aqueous phase re-extracted once with an equal volume of NETS-saturated phenol/chloroform reagent (chloroform/isoamyl alcohol 24:1), 1:1, then once with an equal volume of chloroform reagent. Nucleic acids were precipitated by incubating overnight at –20 °C with 2.5 vol. cold ethanol. After centrifugation, the precipitate was redissolved in a small volume of 10 mM Tris, 1 mM EDTA solution pH 8 (TE); 3 vol. 4 M LiCl solution were added and the precipitate allowed to form for 4 h on ice. After recovery by centrifugation, this precipitate was redissolved and ethanol-precipitated again. The final precipitate (total RNA) was dissolved in water and stored at –70 °C. Total RNA (50 μg) was used in 100 μl of reticulocyte lysate⁶ containing 2 μCi μl^{-1} ^3H -leucine. The mixture was incubated at 30 °C for 90 min. Canine pancreatic microsomes¹⁹, where used, were included at 1 μl per 10 μl of lysate. The nascent form of Con A was immunoprecipitated by standard methods²⁰. Unlabelled Con A was used as a carrier during CNBr cleavage of the three radiolabelled forms. Lane 2 shows the pattern of cleavage fragments from this carrier.



produced from precursor polypeptides like 37CA. The ligation to form mature Con A occurs between residues 118 and 119, and because this link appeared to sequence normally on Edman degradation⁵, it seems to be a peptide bond. The two fragments found in mature Con A (Fig. 1) have been shown to correspond to two halves of Con A, residues 1–119 and 119–237 (ref. 4); these may well be unligated as opposed to proteolytically produced from mature Con A. The RNA sequence coding for the 4-residue N-terminal extension found in Con A from immature seeds lies within a sequence coding for 15 amino acid residues in the middle of the nascent polypeptide; none of this amino acid sequence has been found in mature Con A. The initiation codon is positioned to give a 29-amino acid signal sequence; a signal sequence of the same size has been found and partially sequenced in favin synthesized *in vitro* (favin is the lectin extracted from *Vicia faba*)⁷. There is some sequence homology between the two signal sequences.

All the points at which cleavage must occur to produce the mature Con A molecule (excluding the removal of the signal sequence) have an Asn residue to the N-terminal side of the point of cleavage, a guide to the specificity of the modifying enzyme(s).

The widely found sequence AAUAAA, indicating a polyadenylation site⁸, is not found at the 3' end of the transcribed message, although a similar sequence, (U)AUAAA, is.

Typically for legume storage proteins, the signal sequence is rich in sulphur-containing amino acids compared with the remainder of the molecule⁹. There is also a small C-terminal extension, as found in other legume seed storage proteins¹⁰. The function of the last two features remains unclear.

The precursor of Con A shows direct amino acid sequence homology with other legume lectins. The transposition and ligation occur around a position corresponding to known fragments. This processing could apparently occur without the need for a large change in the three-dimensional structure of Con A⁵. We believe that the sequence presented here does not represent DNA complementary to the mRNA for a precursor to a polypeptide closely related ($\pm 95\%$ amino acid homology) to Con A; this would be inconsistent with the kinetics of polypeptide labelling *in vivo*, and there is only one gene detectable at 85% stringency on genomic Southern blots (to be published elsewhere). Rearrangements within the genome are not necessary to explain the circular amino acid sequence homology observed in legume lectins, although a polypeptide ligase is required. One of the most puzzling features is the function of the ligation, as the fragmented monomer appears to have the same lectin activity as the intact monomer.

Some of this work was carried out while D.M.C. was employed as a Research Assistant on a SERC research grant to D.E.H., and we thank the Nuffield Foundation for a Small Grant Award

A

TCAGAAATGTAGCAACGACACTACTAGTGAAGTGTGAATATCAATAG
1

MAISKSSLSFLPI
GTATACCACCATGCCATCTCAAGAAATCTCCCTGTCTCTCTATAT
51

FTFTITMFLMVVNKVSBS
TTACBTTCATCACCATTCTCAATGTAGTGAACAGGTAGTTCATCA
101

THETNALHFMFNQFSKD
ACACATGAGCAAAATGCACCTCCATTTTCATGTTCACCAATTTAGCAAGA
151

QKDLILQGDAATGTGTEG
TCAGAGGATTTGATCCTTCAAGGTGACGCCCAACAGGAACAGAGGA
201

NLRLTRVSSNGSBPQBS
ACTTGAGACTCAGAGGTGTCAAGTAATGAGAGTCCACAGGAACAGT
251

VGRALFYAPVHIWESSA
GTGGCCGGGCTTGTCTATGCCCAAGTCCACATTTGGAAAGTTTGC
301

VVASFEATFTFLIKSP
TGTGTGTCACGCTTGAAGTACCTTACATTTCTCATAAATACCCG
351

DSHPADGIAFFISNIDS
ACTCTACCCAGCTGATGGAATGTCTTCTTCATTTCAATATTGACAGT
401

SIPSGSTGRLLGLFPDA
TCAATCCCTAGTGTGTCACAGGCTCTTGGAGTCTTCCCTGATGC
451

NVIRNSTTIDFNAYN
AAATGTATCAGAAATCCACTACTATTGATTTCAACGCTGCTACAAT
501

ADTIVAVELDTYPNTDI
CCGATACATTGTGTCTGTGATGATGATCTTATCCCAATCTGATAT
551

BDPSYPHIGIDIKSVRS
GGAGATCCAAGTATTCACACATCBTATCTGATTAATCTGTCTGCTC
601

KKTAKWNMQNGKVGTA
CAAGAGAGCCGCAAGTGAACATGCAAAATGGAAGGTAGGCACTGCAC
651

HIINYNSVDKRLSAVVS
ACATCATCTATAACTCTGTGTATGAGAGACTAAGTGTCTGTCTTAT
701

PNADSBATVSYDVLNDV
CCTAACGCTGACTGCTGCTCTTACGAGCTTGACCTGACATGT
751

LPEWVRVGLSASTGLY
CCTTCTGATGAGTGTAGATGTGGCTTCTGCTTCAACCGACTTTACA
801

KETNTILSMFSFTSKLKS
AGAAACCAATACCATCTCTCATGTCTTTTACTTCAAGTTGAGAGC
851

NEIPDIATVV!!!
AATGAGATCCCGACATGTCTACCTGTGTTGATTTCTCATCCATGTACCA
901

TAATACTCTGCTGCTGTGTATCCACTGTTGGCCTAATATGCCATTAA
951

TGAAGCAATATAAGTCTCTTTATATAAAAAAAAAA
1001

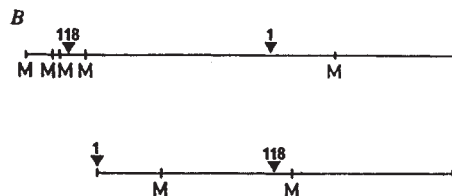


Fig. 4 A, Nucleotide sequence of the cDNA insert in pCE2D11 (see below) in the same sense as the mRNA and the derived amino acid sequence. The numbers below each line correspond to the number of the base in the cDNA; those above the line indicate the number of the amino acid in the published sequence of Con A⁵ (the differences from that sequence are shown also). The signal sequence is indicated by the thick solid line, the amino acids corresponding to residues 1–118 in Con A by the thin solid line, and those corresponding to residues 119–237 by the broken line. B, Positions of methionines (M) in a, nascent *in vitro* Con A including the signal sequence; and b, the mature molecule. The positions of residues 1 and 118 from mature Con A are shown in both molecules

Methods: A, Polyadenylated RNA was purified from total RNA using oligo(dT)–cellulose²¹. Double-stranded cDNA was made by standard methods^{22,23}. After treatment with S₁ nuclease, the cDNA was phenol-extracted and size-fractionated on a 0.6 × 32 cm column of Biogel A15 M. The cDNA was repaired in a reaction containing 50 ng cDNA, 100 ng SmaI-cut pUC8 and 2 U of Klenow fragment of DNA polymerase I in 5 µl of a solution comprising 50 mM Tris pH 7.5, 10 mM MgCl₂, dithiothreitol and 25 µM each of dATP, dTTP, dGTP and dCTP. After 15 min at room temperature, this mixture was diluted to 30 µl with 2 U T4 DNA ligase (Amersham) in 25 µl of the recommended buffer. After overnight ligation at 15 °C, the resultant plasmids were used to transform *Escherichia coli* JM101. Screening was performed by colony hybridization²⁴ using single-stranded DNA complementary to the original poly(A)⁺ RNA. Strongly hybridizing colonies were grouped into homologous sets by colony hybridization using purified cDNA inserts as probes; each set was identified by performing hybrid release translation²¹ on one member. Two clones coding for Con A were sequenced: pCE2B7 (720 base pairs, bp) and pCE2D11 (1,030 bp); they gave the same sequence over the homologous region. The data shown represent the sequence of the longer insert. Five other cDNA clones gave restriction endonuclease patterns consistent with this sequence. No *Eco*RI or *Bam*HI sites were present in the cDNA of pCE2D11. M13 subclones were prepared by digesting 20 µg of *Eco*RI- or *Bam*HI-cut pCE2D11 with 2 U of *Bal*31 (New England Biolabs) in 100 µl of the suggested buffer at 30 °C. Aliquots (10 µl) were taken at 0, 3, 5, 7, 9, 11 and 13 min and added immediately to phenol. After phenol and chloroform extractions and ethanol precipitation, the pooled, shortened inserts were separated from the vector by digestion with *Eco*RI if the initial cut had been with *Bam*HI or vice versa. The reaction mixture was phenol-extracted and the DNA recovered by ethanol precipitation. Depending on the nature of the final restriction enzyme digestion, the DNA was ligated into *Eco*RI- and *Sma*I-digested M13mp8 or *Bam*HI- and *Sma*I-digested M13mp9. The resultant subclones were sequenced by the dideoxy method^{25,26}.

which enabled us to complete it. We particularly thank Drs Mervyn Turner and John Cordingley for help and advice.

Received 6 September; accepted 22 October 1984.

- Hemphery, J. J. & Cunningham, B. A. *Trends biochem. Sci.* **8**, 100–102 (1983).
- Foriers, A., Lebrun, E., van Rapenbusch, R., De Neve, R. & Strosberg, A. D. *J. biol. Chem.* **256**, 5550–5560 (1981).
- Hendrix, R. W. & Casjens, S. R. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1451–1455 (1974).
- Abe, Y., Iwabuchi, M. & Ishii, S.-I. *Biochem. biophys. Res. Commun.* **45**, 1271–1277 (1971).
- Cunningham, B. A., Wang, J. L., Waxdal, M. J. & Edelman, G. M. *J. biol. Chem.* **250**, 1503–1512 (1975).
- Pelham, H. R. B. & Jackson, R. J. *Eur. J. Biochem.* **67**, 247–256 (1976).
- Hemphery, J. J., Mostov, K. E. & Cunningham, B. A. *J. biol. Chem.* **257**, 7903–7909 (1982).
- McDevitt, M. A., Imperiale, M. J., Ali, H. & Nevins, J. R. *Cell* **37**, 993–999 (1984).
- Spencer, D. *Phil. Trans. R. Soc. B304*, 275–285 (1984).
- Boulter, D. *Phil. Trans. R. Soc. B304*, 323–332 (1984).

- Laemmli, U. K. *Nature* **227**, 680–684 (1970).
- Bridgen, J. *FEBS Lett.* **50**, 159–162 (1975).
- Laursen, R. A. *Meth. Enzym.* **47**, 227–288 (1977).
- Brauer, A. W., Margolies, M. M. & Haken, E. *Biochemistry* **14**, 3029–3035 (1975).
- Tarr, G. E., Beecher, J. F., Bell, M. & McKean, D. J. *Analyt. Biochem.* **84**, 622–627 (1978).
- Renart, J., Reiser, J. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3116–3120 (1979).
- Bailey, D. S., Deluca, V., Durr, M., Verma, D. P. S. & MacLachlan, G. A. *Pl. Physiol.* **66**, 1113–1118 (1980).
- Walker, J. E. *et al. Eur. J. Biochem.* **123**, 253–260 (1982).
- Scheele, G., Dobberstein, B. & Blobel, G. *Eur. J. Biochem.* **82**, 593–599 (1978).
- Goldman, B. M. & Blobel, G. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5066–5070 (1978).
- Cordingley, J. S., Taylor, D. W., Dunne, D. W. & Butterworth, A. E. *Gene* **26**, 25–39 (1978).
- Buell, G. N., Wickens, M. P., Payvar, F. & Schimke, R. T. *J. biol. Chem.* **253**, 2471–2482 (1978).
- Wickens, P. C., Buell, G. N. & Schimke, R. T. *J. biol. Chem.* **253**, 2483–2495 (1978).
- Grunstein, M. & Wallis, J. *Meth. Enzym.* **68**, 379–388 (1979).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
- Sanger, F. & Coulson, A. R. *FEBS Lett.* **87**, 107–110 (1978).
- Murashige, T. & Skoog, F. *Physiol. Pl.* **15**, 473–497 (1962).

Synthesis of brome mosaic virus subgenomic RNA *in vitro* by internal initiation on (–)-sense genomic RNA

W. A. Miller, T. W. Dreher* & T. C. Hall*

Biophysics Laboratory, 1525 Linden Drive, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA

The genomes of many (+)-stranded RNA viruses, including plant viruses and alphaviruses, consist of polycistronic RNAs whose internal genes are expressed via subgenomic messenger RNAs. The mechanism(s) by which these subgenomic mRNAs arise are poorly understood. Based on indirect evidence, three models have been proposed: (1) internal initiation by the replicase on the (–)-strand of genomic RNA^{1,2}, (2) premature termination during (–)-strand synthesis, followed by independent replication of the subgenomic RNA³ and (3) processing by nuclease cleavage of genome-length RNA⁴. Using an RNA-dependent RNA polymerase (replicase) preparation from barley leaves infected with brome mosaic virus (BMV)^{5–7} to synthesize the viral subgenomic RNA *in vitro*, we now provide evidence that subgenomic RNA arises by internal initiation on the (–)-strand of genomic RNA. We believe that this also represents the first *in vitro* demonstration of a replicase from a eukaryotic (+)-stranded RNA virus capable of initiating synthesis of (+)-sense RNA.

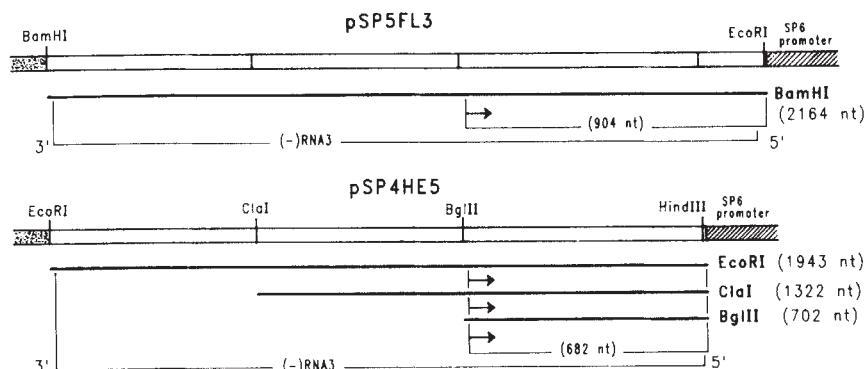
Subgenomic RNA4 of BMV serves as the coat protein mRNA⁸. Its entire sequence is present at the 3' end of dicistronic genomic RNA3, from which it is generated⁹. Minus-stranded copies of both of these RNAs can be synthesized by the template-dependent BMV RNA-specific replicase extract^{5,7}. However, because of the difficulty in obtaining (–)-stranded RNA from infected tissue, the ability of this replicase to synthesize (+)-stranded RNA from (–)-stranded template has not been studied. By using an *in vitro* transcription system described recently for the synthesis of active (+)-sense templates from cloned cDNA¹⁰, we have now generated (–)-sense BMV RNAs. Two transcriptional plasmids containing (–)-BMV RNA3 sequences were constructed (Fig. 1). One (pSP5FL3) contained the whole sequence, the other (pSP4HE5) lacked 200 bases at the 5' end. The transcribed regions of both clones also contained short heterologous vector sequences at both ends. RNAs with various 3' ends were produced by run-off transcription from plasmids linearized at unique restriction sites (Fig. 1).

The BMV replicase preparation was able to use (+)- and (–)-stranded RNAs as templates with comparable activity (Fig. 2). Whereas virion RNAs 3 and 4 gave rise only to full-length products (Fig. 2, lanes 1–4), the predominant product from (–)RNA3 was the size of subgenomic RNA4. The products obtained when the *Bam*HI and *Eco*RI transcripts served as templates (Fig. 2, lanes 5–8) were consistent with the occurrence of internal initiation at the 3' end of the (–)RNA4 sequence (expected sizes 904 and 682 bases respectively). The same products predominated when (–)RNA3 transcripts lacking major segments 3' to the RNA4 region served as templates (*Cla*I and *Bgl*II transcripts; Fig. 2, lanes 9–12). Thus, it appears that the recognition region for internal initiation by BMV replicase on (–)RNA3 does not extend far beyond the (–)RNA4 sequence, as the *Bgl*II transcript extends only 20 bases into this region. The lack of RNA4-sized product from the (+)-sense RNA3 template precluded the possibility that subgenomic RNA arises by premature termination. The absence of significant levels of full-length product from (–)RNA3 may be due to the presence of heterologous bases interfering with normal recognition of the 3' end by the replicase. However, the small amount of full-length product that was synthesized even in the absence of the native 3' end (for example, from the *Cla*I and *Bgl*II transcripts) may be caused by sliding of the polymerase from the internal RNA4 initiation site to the 3' end of the template before initiation. Nonspecific end-to-end copying of RNAs is uncharacteristic of this replicase extract, which has been shown to be highly discriminatory against a wide range of RNAs^{5,7}.

Typical products synthesized by BMV replicase from virion RNAs are double-stranded ribonuclease-resistant structures^{5,7}. Products synthesized from (–)RNA templates were also ribonuclease resistant (Fig. 3, lanes 1–4), suggesting that (+)-stranded RNA was made. Hybridization of replicase products to strand-specific M13 clones of BMV sequences produced more definitive evidence for this. Following removal of the large excess of unused template RNA by treatment with RNase, single-stranded DNA containing nearly full-length RNA4 sequences in the (–)- or (+)-sense orientations (M13 clones S2T1 and S2, respectively)¹¹ was hybridized to denatured replicase products. Hybrids resistant to S₁ nuclease were analysed electrophoretically. All the label incorporated into the product from the (–)-sense (*Eco*RI) transcript was protected from nuclease digestion by hybridization to S2T1, but not by hybridization to S2 (Fig. 3, lanes 6, 7), indicating that only (+)-sense RNA4 was synthesized. In contrast, the product from (+)RNA4 was shown to be (–)-sense by hybridization to S2 (Fig. 3, lanes 8, 9). The protected fragment of the replicase product from the (–)-stranded template was the same size as, but in the opposite

* To whom correspondence should be addressed; present address: Department of Biology, Texas A & M University, College Station, Texas 77843, USA.

Fig. 1 Transcriptional plasmids used for synthesis of (–)-BMV RNA3. A full-length complementary DNA clone of BMV RNA3⁹ was subcloned into the transcriptional vector pSP65 (ref. 19) using *Sma*I and *Eco*RI sites present outside the BMV sequence. Transcription of the resultant plasmid (pSP5FL3), after linearization with *Bam*HI, produced RNA containing the entire (–)-BMV RNA3 sequence, with stretches of 28 and 19 heterologous bases (introduced during cDNA cloning) at the 5' and 3' ends, respectively. A second transcriptional plasmid (pSP4HE5) was constructed by subcloning an *Eco*RI–*Hind*III segment of the full-length cDNA into the vector pSP64. Transcription of this plasmid after linearization with *Eco*RI produced (–)RNA3 lacking 200 bases at the 5' end, and containing 6 and 20 heterologous bases at the 5' and 3' ends, respectively. The figure shows the transcribed regions of these plasmids (open bar) and the unique restriction sites at which they were linearized before run-off transcription to yield the transcripts indicated (bold lines beneath plasmids). The boundary of the (–)RNA3 sequence is shown. Arrows indicate the direction of polymerization by the replicase and expected internal initiation site at the 3' end of the (–)RNA4 sequence. The expected sizes of the products of internal initiation, including vector sequences at the 5' ends, are indicated below the transcripts. For brevity, each transcript is referred to in the text by the restriction site at which the plasmid was linearized; only these RNAs are called transcripts. Replicase products are referred to as products.



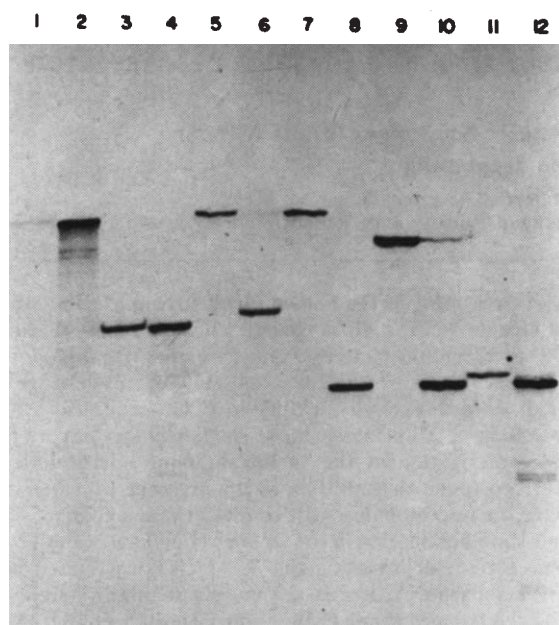


Fig. 2 Products of BMV replicase using (–)RNA3 transcripts as templates. Viral RNAs and transcripts were ^{32}P -labelled as markers (odd-numbered lanes), but were unlabelled or ^3H -labelled when used as templates for synthesis of ^{32}P -labelled products of replicase (even-numbered lanes). RNAs were analysed by electrophoresis on 4% polyacrylamide, 7.7 M urea denaturing gels²⁰ followed by autoradiography: BMV RNA3 (lane 1) and product (lane 2), BMV RNA4 (lane 3) and product (lane 4), *Bam*HI transcript (lane 5) and product (lane 6), *Eco*RI transcript (lane 7) and product (lane 8), *Clal* transcript (lane 9) and product (lane 10), *Bgl*II transcript (lane 11) and product (lane 12).

Methods: Transcripts were prepared from transcriptional plasmids linearized by digestion with the indicated restriction enzyme (Fig. 1), using SP6 RNA polymerase (Promega-Biotec), as described previously¹⁰, except that incubations were at 40°C for 2–3 h. Marker transcripts were labelled with [α - ^{32}P]UTP, and transcripts used as replicase templates were labelled with ^3H -UTP and produced in 100- μl reactions containing 10 μg plasmid DNA and 20 U SP6 RNA polymerase. Following extraction with phenol-chloroform, two cycles of ethanol precipitation with 2.5 M ammonium acetate were used to remove unreacted NTPs. Because the presence of plasmid DNA did not affect results, transcripts were not treated with DNase (except in Fig. 4). ^3H -labelled transcripts were analysed by electrophoresis and fluorography before use as templates (data not shown). BMV replicase reactions (50 μl final volume) containing 30 μl micrococcal nuclease-treated replicase⁷, 5 μM [α - ^{32}P]UTP (92 Ci mmol⁻¹), 1–2 μg ^3H -labelled transcript or unlabelled viral RNA template, 50 mM Tris pH 8.0 and 10 mM Mg-acetate⁷ were incubated at 30°C for 30 min and the reaction was terminated by phenol-chloroform extraction. Virion RNAs used as markers were labelled with ^{32}P -pCp and RNA ligase²¹.

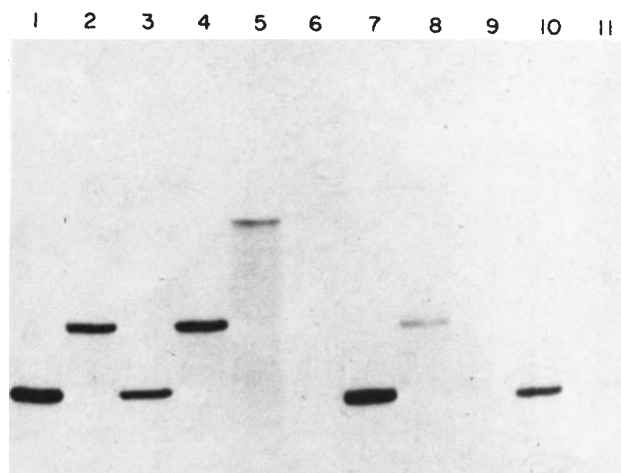


Fig. 3 Strandedness of replicase products. Hybridization of products to strand-specific M13 clones to demonstrate complementary strand synthesis by the replicase. The lanes show denaturing gel electrophoresis of ^{32}P -labelled replicase products from ^3H -*Eco*RI transcript (1) and RNA4 (2), the same products after RNase treatment (3 and 4, respectively), ^{32}P -labelled *Eco*RI transcript (5), *S*₁ nuclease-treated products of replicase, using ^3H -*Eco*RI transcript as template, after hybridization to S2 (6) or S2T1 (7), *S*₁ nuclease-treated products, using RNA4 as template, after hybridization to S2 (8) or S2T1 (9), *S*₁ nuclease-treated ^{32}P -*Eco*RI transcript after hybridization to S2 (10) or S2T1 (11).

Methods: Single-stranded cDNA clones, containing all but 12 bases at the 5' end of the (+)-strand of the BMV RNA4 sequence were constructed in the (+) (S2) and (–) (S2T1) orientations in phages M13mp9 and M13mp8, respectively¹¹. Products of replicase using 5 μg of RNA4 or ^3H -*Eco*RI transcript as templates were synthesized and purified as in Fig. 2, but labelled at a higher specific activity (180 Ci mmol⁻¹, 10 μM [α - ^{32}P]UTP). The products were treated with a mixture of ribonucleases A and T1 in high salt, followed by proteinase K digestion⁵. After phenol-chloroform extraction and ethanol precipitation with 2 μg transfer RNA as carrier, each reaction product, as well as a sample of ^{32}P -*Eco*RI transcript, was divided into two tubes containing 7.5 μg of S2 or S2T1 DNA. These mixtures were then boiled for 2 min, cooled on ice, adjusted to a final volume of 20 μl containing 50 mM PIPES pH 6.4, 0.3 M NaCl and hybridized at 50°C for 45 min¹¹. To this mixture was added 1/10 volume of 0.5 M sodium acetate pH 4.0, 0.5 M NaCl, 60 mM ZnCl₂ and 10 units of *S*₁ nuclease (Pharmacia P-L). After incubation (37°C, 20 min), the reaction was stopped with 1 μl 200 mM EDTA, phenol-chloroform extracted and ethanol precipitated in the presence of 2 μg tRNA. One-third of each final sample of RNA-DNA hybrid was analysed on a 4% polyacrylamide, 7.7 M urea denaturing gel.

orientation to, the protected fragment of the transcript itself (Fig. 3, lanes 10, 11). This confirms that the replicase product contained the complete, contiguous RNA4 sequence complementary to that present on the template.

To establish that the RNA4-sized products were arising by internal initiation, products were synthesized in the presence of each [γ - ^{32}P]-labelled nucleoside triphosphate (NTP). A product identical in size to that synthesized in the presence of [α - ^{32}P]UTP was observed when [γ - ^{32}P]GTP was the labelled substrate (Fig. 4). No products were visible when other [γ - ^{32}P]NTPs were used, confirming that internal initiation was occurring and that G was the first nucleotide of the RNA4-sized product, as expected from the known sequence¹². The low level of radioactivity incorporated into this end-labelled product prevented sequencing. We have also shown that BMV replicase initiates (–)-strand synthesis on virion RNA with a G (J. J. Bujarski, W.A.M., T.W.D. and T.C.H., manuscript in preparation).

These experiments show unequivocally that an enzyme preparation from BMV-infected barley leaves is able to utilize both (+) and (–)-stranded BMV RNAs and that the subgenomic RNA4 is generated by internal initiation from (–)RNA3. The question of independent replication of RNA4 was not examined, but *in vivo* studies have indicated that BMV RNA4 present in the inoculum remains unexpressed and thus presumably unreplicated⁸. This suggests that (–)RNA4 lacks a replicase recognition signal at its 3' end, a possibility being tested. Computer-generated secondary structures¹³ (not shown) indicate that a stable hairpin, with the RNA4 initiation site located at the start of the loop, can form when as few as 20 bases of (–)RNA3 outside the (–)RNA4 sequence are present (*Bgl*II transcript). No such structure can form at the 3' end of (–)-BMV RNA4, which is unusually devoid of secondary structure. We propose that this hairpin stem is part of the recognition signal for initiation of subgenomic RNA synthesis. As this structure bears no obvious

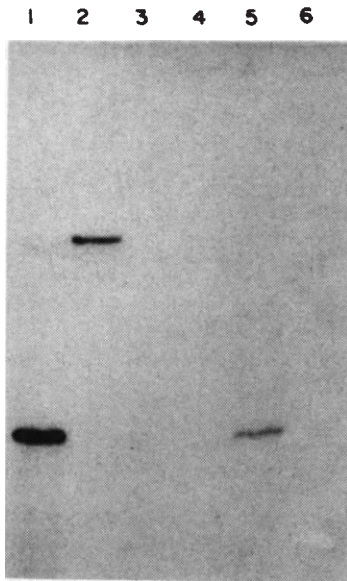


Fig. 4 Incorporation of [γ - 32 P]GTP to demonstrate initiation of RNA4 synthesis by replicase. Denaturing gel electrophoresis of replicase products synthesized using 3 H-*Eco*RI transcript as template. Product was synthesized using [α - 32 P]UTP as substrate (lane 1), marker 32 P-labelled *Eco*RI transcript (lane 2), replicase products synthesized using [γ - 32 P]ATP (lane 3), [γ - 32 P]CTP (lane 4), [γ - 32 P]GTP (lane 5) and [γ - 32 P]UTP (lane 6) as labelled substrates. **Methods:** NTPs labelled at their γ -phosphates were prepared and analysed using a Gamma Prep kit from Promega Biotec. Replicase reactions (50 μ l) contained 30 μ l BMV replicase, 200–400 μ Ci of one of the four [γ - 32 P]NTPs, 0.67 mM of each of the three unlabelled NTPs, and 2.3 μ g DNase-treated 3 H-*Eco*RI transcript (Fig. 2). After 5 min at 30 $^{\circ}$ C, reactions were chased for 15 min with the unlabelled form of the 32 P-labelled NTP (final concentration 0.4 mM). Electrophoresis was through a 4% polyacrylamide, 7.7 M urea denaturing gel.

similarity to that recognized specifically by BMV replicase at the 3' end of the (+)RNAs¹¹, it is possible that different functional proteins present in the membrane-bound enzyme preparation are responsible for the activities observed. The involvement of different viral proteins in the synthesis of genomic and subgenomic RNAs of plant viruses^{14,15} and alphaviruses¹⁶ has been demonstrated from mutant studies. The homologies detected recently between the non-structural proteins of the alphaviruses and several plant viruses, including BMV^{17,18}, suggest that these viruses have similar replication mechanisms, including the mode of subgenomic RNA generation.

We thank P. Kaesberg for providing laboratory space, P. Ahlquist for the BMV cDNA clones and C. Harkin and J. Weiland for technical assistance. This work was funded by NIH grant AI11572.

Received 29 August; accepted 12 October 1984.

- Kennedy, S. I. T. in *The Togaviruses* (ed. Schlesinger, R. W.) 351–369 (Academic, New York, 1980).
- Nassuth, A., Alblas, F. & Bol, J. F. *J. gen. Virol.* **53**, 207–214 (1981).
- Goelet, P. & Karn, J. *J. molec. Biol.* **154**, 541–550 (1982).
- Gonda, T. J. & Symons, R. H. *J. gen. Virol.* **45**, 723–736 (1979).
- Hardy, S. F., German, T. L., Loesch-Fries, S. & Hall, T. C. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4956–4960 (1979).
- Bujarski, J. J., Hardy, S. F., Miller, W. A. & Hall, T. C. *Virology* **119**, 465–473 (1982).
- Miller, W. A. & Hall, T. C. *Virology* **125**, 236–241 (1983).
- Lane, L. C. *Adv. Virus Res.* **19**, 151–220 (1974).
- Ahlquist, P., Luckow, V. & Kaesberg, P. *J. molec. Biol.* **153**, 23–38 (1981).
- Dreher, T. W., Bujarski, J. J. & Hall, T. C. *Nature* **311**, 171–175 (1984).
- Ahlquist, P., Bujarski, J. J., Kaesberg, P. & Hall, T. C. *Pl. molec. Biol.* **3**, 37–44 (1984).
- Dasgupta, R. & Kaesberg, P. *Nucleic Acids Res.* **10**, 703–713 (1982).
- Zuker, M. & Stiegler, P. *Nucleic Acids Res.* **9**, 133–148 (1981).
- Dawson, W. O. *Virology* **115**, 130–136 (1981).
- Nassuth, A. & Bol, J. F. *Virology* **124**, 75–85 (1983).
- Strauss, E. G. & Strauss, J. H. *Curr. Topics Microbiol. Immun.* **105**, 1–97 (1983).
- Haseloff, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 4358–4362 (1984).
- Ahlquist, P., Strauss, E. G., Rice, C. M., Strauss, J. H. & Zimmern, D. *J. Virol.* (submitted).
- Melton, D. A. *et al. Nucleic Acids Res.* **12**, 7035–7056 (1984).
- Sanger, F. & Coulson, A. R. *FEBS Lett.* **87**, 107–110 (1978).
- England, T. E., Bruce, A. G. & Uhlenbeck, O. C. *Meth. Enzym.* **65**, 65–74 (1980).

Reversible association of myosin with the platelet cytoskeleton

Vivianne T. Nachmias, Joshua Kavalier & Sam Jacobowitz

Department of Anatomy, School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

Platelets circulating in the human blood stream are smooth disk-shaped structures. The disks change within seconds of exposure to ADP or thrombin to irregular spheres bearing filopodia and pseudopodia. It is well-established that platelets also change shape (although more slowly) when chilled to 5 $^{\circ}$ C^{1–5} and revert to disks on rewarming^{1,3}. This cold-induced shape change may be due to the depolymerization of the submembranous microtubule ring. However, we found that chilling in the presence of Taxol, which stabilizes the microtubules, still results in shape change. Chilled platelets show an increase in the amount of myosin in the Triton-X insoluble residue or 'cytoskeleton'^{6–9} which is correlated in time both with phosphorylation of the myosin regulatory light chain and with the induced shape change. We suggest here that the slow cold-induced change from disks to spheres is due primarily to a gradual activation of myosin.

The amount of myosin heavy chain in the platelet cytoskeleton increases 8–13 times above control levels after 60 minutes in the cold (Fig. 1). On rewarming to 37 $^{\circ}$ C, myosin dissociates from the cytoskeleton. As electron microscopy demonstrates the disappearance of the microtubule ring and an increase in the number of visible microfilaments in chilled platelets³, myosin solubility may be altered by a change in state of actin or tubulin^{10,11}. However, addition of rabbit muscle F-actin at 0.1–

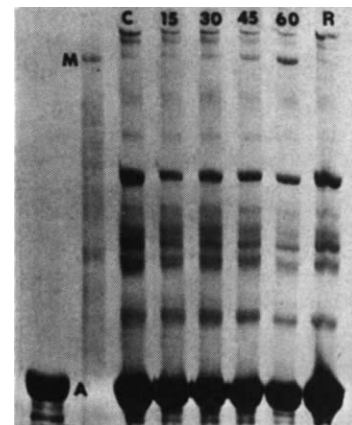


Fig. 1 SDS gel electrophoresis showing myosin heavy chain in the cytoskeleton as a function of time of chilling of platelets. Platelets were filtered at 37 $^{\circ}$ C from platelet rich plasma through a column of Sepharose 2B equilibrated and eluted with 0.1 M KCl, 10 mM imidazole, 10 mM EGTA pH 6.8 and divided into aliquots. These were kept for one h at 37 $^{\circ}$ C (lane 3, C) or in ice at 2 $^{\circ}$ C for 15, 30, 45 and 60 min (lanes 4–7) or in ice for 60 min followed by 37 $^{\circ}$ C for another 60 min (lane 8, R). The platelets were then lysed by adding 0.10 vol 10% Triton-X 100 and 0.01 vol 100 mM phenylmethylsulphonyl fluoride. The lysate was immediately put on ice for 10 min and centrifuged at 1,400g for 10 min. The insoluble fractions, the 'cytoskeletons', were then washed twice with 5 ml vols of the above buffer, ice cold. These precipitates were taken up into 40 μ l of 2 $\times$ Laemmli buffer and were electrophoresed on a 7.5% SDS gel using the same method²⁷. Lane 1, actin standard; lane 2, myosin standard; arrow, position of the myosin heavy-chain band. Note that by 30 min there is a detectable increase in the amount of myosin in the cytoskeletons, and by 1 h the increase is marked. After an hour of rewarming, the myosin band has decreased again, almost to the original low level. The remaining components of the cytoskeleton appear constant as a function of time of chilling.

0.5 mg ml⁻¹ to suspensions of resting platelets immediately before lysis did not cause myosin heavy chain to complex with the cytoskeletons although there was an increase in the amount of precipitated actin. Conversely, addition of colchicine to platelets in the cold before rewarming did not inhibit the return of myosin to the supernatant fraction.

We therefore sought a change in state of the myosin. Chilling radiolabelled platelets induces phosphorylation of a band on SDS gels with the mobility of myosin light chain (Fig. 2, lanes 1-4). ³²P is present also in the light chain of myosin when the molecule is extracted subsequently from the cytoskeletons with ATP and pyrophosphate (Fig. 2, lanes 5-8).

Densitometry of myosin light chain phosphorylation in whole platelet lysates and in cytoskeletons demonstrates that phosphorylation of light chains increases with time to 9-15 times above control levels after 60 minutes in the cold, in agreement with the heavy chain results (Table 1). Phosphorylation is reversed by warming for one hour, as with heavy chain precipitation.

Part of the signal for platelet and megakaryocyte activation may involve increased permeability to sodium¹²⁻¹⁶. Furthermore, platelet shape change can be produced by monovalent ionophores¹⁷. We therefore ask whether cold-induced myosin phosphorylation requires external sodium, and demonstrate conclusively that there is no inhibition of the light chain phosphorylation when sodium is removed (Fig. 3). Densitometric comparisons show that phosphorylation due to cold or to thrombin is not decreased by sodium removal. The level of phosphorylation after chilling is 30-50% of that caused by 0.1 U ml⁻¹ thrombin. By comparison with published values^{6,18} for thrombin, 27-45% of the myosin is phosphorylated by exposing platelets to cold for one hour.

Because myosin light chain kinase is a calcium- and calmodulin-dependent enzyme^{19,20}, we examined the effects of inhibitors of such enzymes on cold-induced phosphorylation. We found variable inhibitions of up to 50% of the coprecipitation of the heavy chain and up to 40% of light chain phosphorylation. However, high doses were needed (60 mM) and the inhibition was not as reliable nor as complete as with thrombin-induced phosphorylation in parallel experiments and as reported in the literature⁶.

The correlation of myosin light chain phosphorylation with myosin in the cytoskeleton is unlikely to be caused by the precipitation of myosin alone, as the centrifugal forces in the experiments (1,400g for 10 minutes or 12,000g for 2 minutes) are not high enough. Myosin associated with actin filaments is likely to be precipitated more efficiently due to increased interaction of myosin molecules with each other. This could be caused by changes in configuration of the myosin molecule after phosphorylation²¹⁻²³ resulting in enhanced filament formation²⁰. Increased association with actin could also occur as postulated for similar thrombin-induced changes⁶.

It is surprising that phosphorylation is induced by cold, which is often used to stop reactions, and we next questioned whether this phenomenon is related to cold-induced shape change. It

has been reported that myosin light chain phosphorylation seen after chilling platelets for 10 min in the cold required external calcium, whereas shape change did not, suggesting that these responses were uncoupled. Our experiments are carried out in 10 mM EGTA, so that the light chain phosphorylation cannot

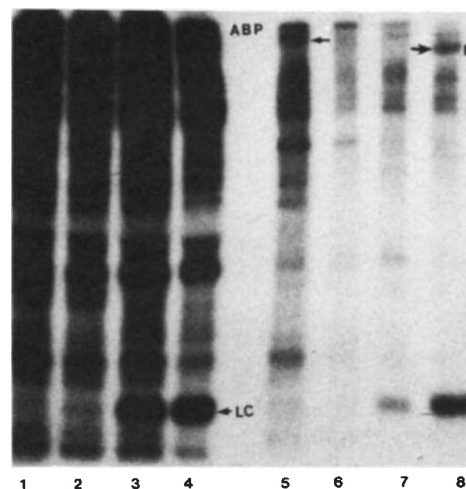


Fig. 2 Lanes 1-4, phosphorylation of the proteins in cytoskeletons; lanes 5-8, phosphorylation in the supernatant after ATP/pyrophosphate extraction of the cytoskeletons.

Methods: A concentrated suspension of resting platelets was collected from platelet rich plasma by centrifugation onto a step gradient of 40% bovine serum albumin in a buffer of 140 mM NaCl, 4.4 mM KCl, 10 mM EGTA, 10 mM imidazole, pH 7.0 (buffer A) and was incubated with 1 mCi ml⁻¹ [³²P] (NEN) at 37 °C for 1 h. The platelets were then gel filtered in buffer A as Fig. 1 and kept at 37 °C (lanes 1-2 and 6-7) or put in ice for 1 h (lanes 3-4 and 7-8). An aliquot of platelets was lysed as above and after 10 min in ice the cytoskeletons were spun down in an Eppendorf microfuge, washed twice in the same way and dissolved in Laemmli buffer (lanes 1-4); the remaining part of the suspension was lysed, precipitated and washed in the same way but the precipitate was then extracted with a buffer containing 2 mM ATP, 1 mM pyrophosphate, 4 mM MgCl₂, 5 mM dithiothreitol and the supernatant dissolved in 5 µl 4× Laemmli buffer. All samples contained 25 mM sodium fluoride added directly after lysis and again to each of the two wash solutions used for the cytoskeletons. Lanes 2, 4, 6 and 8 were samples to which EDTA was also added at the time of lysis to give 10 mM final concentration. Electrophoresis was on 13% gels. Note that a band with the mobility of myosin light chain, as determined by a standard on the Coomassie blue stained gel (LC in lane 5) is phosphorylated strongly by the cold treatment in the cytoskeletal samples (lanes 3 and 4). Phosphorylated myosin light chain is present in the supernatants from these samples after the ATP/pyrophosphate extraction (lanes 7 and 8). The addition of EDTA at the time of lysis and in the wash steps resulted in the appearance of phosphorylation in the position of myosin heavy chain (arrow, lane 8). The 37° extract contains no phosphorylated light chain but does contain a phosphorylated band at the position of actin binding protein with a mobility less than that of myosin as previously observed³.

Table 1 Relative phosphorylation of myosin light chain in whole platelet lysates and cytoskeletal fraction of platelets incubated at 37 and 4 °C

Condition	Expt 1 Autoradiography of whole platelet lysates		Expt 1 Autoradiography of cytoskeletons		Expt 2 Autoradiography of cytoskeletons	
	Area (arbitrary units)	% of maximum	Area (arbitrary units)	% of maximum	Area (arbitrary units)	% of maximum
37 °C 60 min	3.0	28	0.7	11	0.3	6
4 °C 20 min	9.2	87	4.0	66	3.2	62
4 °C 40 min	9.1	86	3.8	62	4.3	83
4 °C 60 min	10.6	100	6.1	100	5.2	100
4 °C 60 min + 37 °C 60 min	6.0	57	0.9	15	0.4	8

Densities in the whole platelet lysates reflect less closely the single contribution of myosin light chain because of other phosphorylated products overlapping the light chain region. Note the 9-fold increase in myosin light chain region density of cytoskeletons after 60 min at 4 °C. Measurements were made in duplicate with a Canalco Model K densitometer and areas determined with a Numonics 1,250 planimeter.

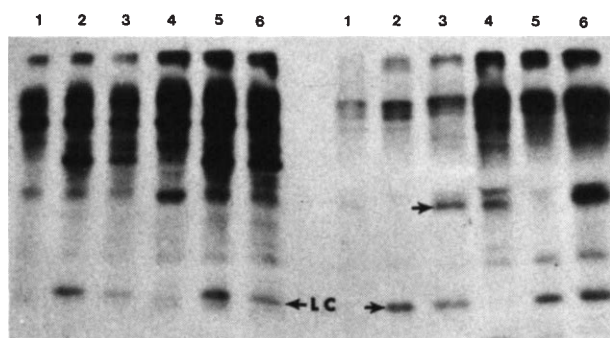


Fig. 3 Phosphorylation of the myosin light chain is unaffected by replacement of sodium chloride with choline chloride and sodium channel blockers. Platelets were preincubated with labelled phosphate as Fig. 2 and then gel filtered either in buffer A, or in buffer A with the sodium chloride replaced with choline chloride plus 1 mM amiloride. To the experimental samples, 2 μ M tetrodotoxin was added also for 10 min before adding thrombin or placing the platelets at 5 °C for 1 h. Left panel, lanes 1–3, whole platelet lysates from resting, thrombin treated (0.1 U ml⁻¹) and 60 min chilling; lanes 4–6, the same treatments but with sodium replaced by choline and with the addition of amiloride and tetrodotoxin as above. Right panel, cytoskeletons from the same experiment. Arrows, the positions of myosin light chain and a relative molecular mass M_r 36,000 protein of unknown properties which appears in the cytoskeleton after chilling but not after thrombin treatment. Both thrombin and cold cause increased phosphorylation of a band with mobility of myosin light chain, and also of a M_r 47,000 band which is present in the supernatants but not in the cytoskeletons. Densitometry of this and similar gels was used to derive an estimate of the extent of phosphorylation in cold compared with thrombin phosphorylation.

be due to external calcium. To examine shape change, we scored more than 100 gel-filtered platelets fixed in 2% formaldehyde/2% glutaraldehyde at several time-points after chilling. The use of thick layers of fluid allowed us to see each scored platelet on edge. We distinguished four states: disks without filopodia, disks with filopodia, irregular or distorted disks and fully altered platelets (irregular spheres). Our results show that at 37 °C platelets are 99% discoid, but 70% of them have one or two filopodia. Chilled platelets become progressively more distorted from the disk shape until they are fully altered. After 20 minutes, 81% are distorted and 17% fully shape changed and after 60 minutes 81% are fully altered and 16% distorted.

Many distorted disks but no spheres are present after 10 minutes of chilling. These disks can be considered fully altered if not compared carefully with later stages. Thus, the degree of phosphorylation (Table 1) in our experiments parallels quite well the degree of shape change from disk to fully altered sphere. The protrusion of filopodia, however, does not appear to depend on myosin phosphorylation as suggested by previous work on the *in vitro* formation of actin bundles without myosin⁷. As shape change has been thought to depend on microtubule breakdown, we also tested the effects of adding Taxol to gel filtered platelets at 30 μ M for 10 minutes before chilling for 1 hour. This treatment preserved the microtubule ring as shown in electron micrographs (taken by Dr Fern Tablin), but did not prevent the appearance of irregular spheres by phase or electron microscopy. This retention of cold-induced shape change is in agreement with previous results where Taxol pretreatment seemed to allow some (incomplete) shape change⁴. Differences in rates of cooling may explain this discrepancy; there may be a stabilizing effect of the microtubule coil at earlier time points during the change from disks to distorted disks.

Reversible light chain phosphorylation also occurs at room temperature using the platelet activator arachidonic acid followed by prostacyclin, which increases cyclic AMP levels²⁴. The presence or absence of myosin in the cytoskeletons was correlated with light chain phosphorylation, platelet aggregation and, to a considerable extent, platelet shape. The identification of

heavy chain phosphorylation observed here when EDTA and fluoride are present during the lysis and wash steps (Fig. 2) suggests that this phosphorylation may be related also to myosin activation. Heavy chain phosphorylation has been reported for macrophages²⁵ in mammalian cells, but no physiological function has yet been found. Daniel *et al.*²⁶ have demonstrated that 50–70% of the myosin light chain is phosphorylated in five seconds following ADP addition in parallel with shape change; both these parameters were inhibited in parallel by ATP. This effect occurs without aggregation or secretion in agreement with Daniel *et al.*¹⁸. The correlations observed here and the lack of inhibition by Taxol suggest that the slow induction of shape change by chilling also depends on myosin phosphorylation.

This investigation was supported by grant HL-15835 to the Pennsylvania Muscle Institute. We thank our volunteers for donating platelet samples, the Philadelphia Red Cross for supplying outdated platelets and P. W. Bronson and M. C. Glennon for assistance.

Received 5 March; accepted 19 September 1984.

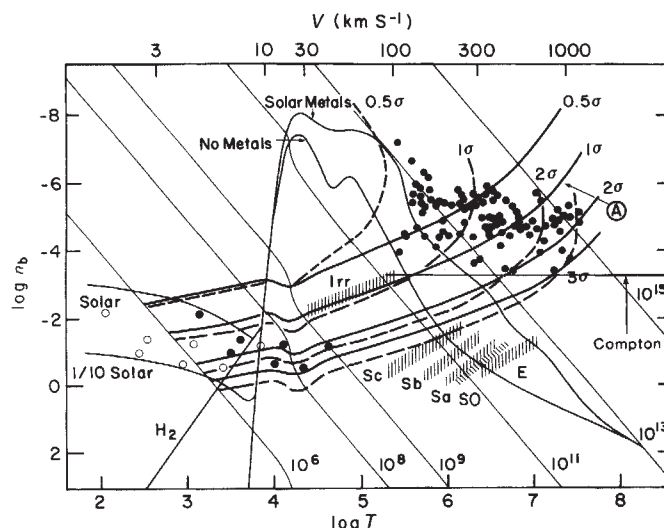
- Behnke, O. *Int. Rev. exp. Path.* **9**, 1–92 (1970).
- Bennett, W. F. & Lynch, G. J. *Cell Biol.* **86**, 280–285 (1980).
- White, J. G. & Krumwiede, M. *Blood* **41**, 823–832 (1973).
- White, J. G. *Am. J. Path.* **108**, 184–195 (1984).
- Zucker, M. B. & Borrelli, J. *Blood* **9**, 602–608 (1954).
- Fox, J. E. B. & Phillips, D. R. *J. Biol. Chem.* **257**, 4120–4126 (1982).
- Gonnella, P. A. & Nachmias, V. T. *J. Cell Biol.* **89**, 746–751 (1981).
- Lucas, R. C., Rosenberg, S., Shafiq, S., Stracher, A. & Lawrence, J. *Protides Biol. Fluids* **26**, 465–470 (1978).
- Nachmias, V. T. *J. Cell Biol.* **86**, 795–802 (1980).
- Hayashi, M., Ohnishi, K. & Hayashi, K. *J. Biochem. (Tokyo)* **87**, 1347–1355 (1980).
- Shimo-Oka, T., Hayashi, M. & Satanabe, Y. *Biochemistry* **21**, 4921–4926 (1980).
- Connolly, T. M. & Limbird, L. E. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5320–5324 (1983).
- Feinberg, H. W. *et al. Biochim. biophys. Acta* **470**, 317–324 (1977).
- Leven, R. M., Mullikin, W. H. & Nachmias, V. T. *J. Cell Biol.* **96**, 1234–1240 (1983).
- Nachmias, V. T. *Sem. Haemat.* **20**, 261–281 (1983).
- Sandler, W. C., Le Breton, G. C. & Feinberg, H. *Biochim. biophys. Acta* **600**, 448–455 (1980).
- Leven, R. M., Gonnella, P. A., Reeber, M. J. & Nachmias, V. T. *Thromb. Haemostasis* **49**, 230–234 (1983).
- Daniel, J. L., Molish, I. R. & Holmsen, H. L. *J. Biol. Chem.* **256**, 7510–7514 (1981).
- Dabrowska, R. & Hartshorne, D. J. *Biochem. biophys. Res. Comm.* **85**, 1352–1359 (1978).
- Hathaway, D. R. & Adelstein, R. S. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1653–1657 (1979).
- Craig, R., Smith, R. & Kendrick-Jones, J. *Nature* **302**, 436–439 (1983).
- Scholey, J. M., Taylor, M. A. & Kendrick-Jones, J. *Nature* **287**, 233–235 (1980).
- Trybus, K. M., Huiatt, T. W. & Lowey, S. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6151–6155 (1981).
- Cox, A. C., Carroll, R. C., White, J. C. & Rao, G. H. R. *J. Cell Biol.* **98**, 8–15 (1984).
- Trotter, J. A. *Biochem. biophys. Res. Comm.* **106**, 1071–1077 (1982).
- Daniel, J. L., Molisch, I. R., Rigmaiden, M. & Stewart, G. J. *J. Biol. Chem.* **259**, 9826–9831 (1984).
- Laemmli, U. K. *Nature* **227**, 680–685 (1970).

Erratum

Formation of galaxies and large-scale structure with cold dark matter

G. R. Blumenthal, S. M. Faber, J. R. Primack & M. J. Rees
Nature **311**, 517–525 (1984)

FIGURE 3 in this review article did not include the curve marked 'Compton'. The correct figure is shown below.



Acidification of soil and water

VAN BREEMEN *et al.*¹ narrowly and arbitrarily define acidification of soil as the 'loss of bases' or 'decrease in acid-neutralizing capacity', the rate of which can be equated with the rate of mineral weathering². However, any definition of soil acidification should include the obvious fact that it also involves the 'accumulation of acids', that is, an increase in 'base-neutralizing capacity'². Although the authors consider the base-neutralizing capacity of water, they do not do so for soils¹. By ignoring soil acidity they conclude that cation exchange reactions play no part in acidification¹, whereas in fact they do³.

The authors' statement that enough is now known to compare external and internal sources of acidity¹ is not supported by the material they present. Estimates of so-called hydrogen-ion budgets are highly data intensive and require an extraordinary amount of effort to develop⁴. Not surprisingly, critical examination of the references cited by the authors and used to develop such budgets, as in Table 2 of ref. 1, shows that for many of the sites the data are seriously incomplete. Even in those cases where the data are complete enough to develop a hydrogen-ion budget, as at Hubbard Brook, the budget is of net transformations that occur throughout a year. Driscoll *et al.* have stated⁵ that "[hydrogen-ion budget] in no way attempts to quantify the gross hydrogen-ion transformations within the ecosystem" and "it is virtually impossible to quantify all of the individual, internal cycling of elements that occur within an ecosystem". Accordingly, total proton sources and sinks have not been categorized, contrary to statements of van Breemen *et al.*¹. Their data are biased towards overestimating the relative extent and rate of acidification by acidic deposition.

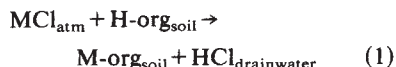
Regions of eastern North America and northern Europe, where there is great concern about acidification, are characterized by the pronounced accumulation of acid organic matter². The role of organic acids is not simply one of cation exchange (or weathering), as asserted by van Breemen *et al.*¹, but also involves complexation and solubilization². Some organic complexes and acids normally soluble in 'clean' environments are immobilized by the strong acids in rain^{2,6}. This compensating shift in internal cycling of elements and acids is not accommodated by hydrogen-ion budgets derived from net transformations.

E. C. KRUG

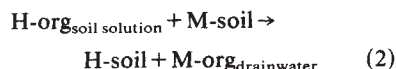
The Connecticut Agricultural
Experiment Station,
New Haven, Connecticut 06504, USA

1. Van Breemen, N., Driscoll, C. T. & Mulder, J. *Nature* **307**, 599-604 (1984).
2. Krug, E. C. & Frink, C. R. *Science* **221**, 520-525 (1983).
3. Overrein, L. N., Seip, H. M. & Tollan, A. (eds) *SNSF Project Res. Rep. FR19/80* (1980).
4. Galloway, J. N., Norton, S. A. & Church, M. R. *Envir. Sci. Technol.* **17**, 541A-545A (1983).
5. Driscoll, C. T. & Likens, G. E. *Tellus* **34**, 283-292 (1982).
6. Krug, E. C. & Isaacson, P. J. *Soil Sci.* **137**, 370-378 (1984).

VAN BREEMEN *ET AL.* REPLY—Accumulation of organic matter does increase the base neutralizing capacity of soils and could be considered as soil acidification. However, changes in organic base neutralizing capacity (which is determined using strong base or a concentrated salt solution) are irrelevant to the proton loading of soils and waters under ambient conditions, where only a very small part of the organic base neutralizing capacity is realized by dissociation of H⁺. Hydrogen-ion budgets¹⁻³ do, in fact, account for two major processes of H⁺ dissociation in organic matter: (1) the displacement of H⁺ from organic matter by atmospherically deposited cations

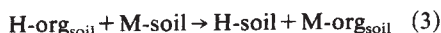


and (2) the dissolution and deprotonation of organic acids followed by leaching



H⁺ budgets show that net H⁺ fluxes, due to organic acids, are zero (equation 1) or relatively small (equation 2), even in podzolic and peat soils⁷. For example, Hubbard Brook Experimental Forest stream water⁴ is dominated by SO₄²⁻ (122.4 µequiv. l⁻¹ out of 166.5 µequiv. l⁻¹ of total anionic charge) largely of atmospheric origin, whereas aqueous acidity (H⁺ + Alⁿ⁺ = 35 µequiv. l⁻¹) far exceeds organic anions (< 2 µequiv. l⁻¹), showing that atmospheric inputs, not organic acids, cause acidification of stream water.

Proton release, at a pH below the reference value (for example, pH 4.5 for forest soils), by complexation of soil metals with organic matter, without drainage export, is ignored in our H⁺ budget, and should indeed be considered as soil acidification:



The only important process of this kind is the accumulation of Al and Fe organic complexes in podzols. Data on Hubbard Brook soils⁵, and titration curves for H⁺-saturated organic matter from podzol-B horizons^{6,7}, reveal that this process contributes little (0.005-0.01 kmol hectare⁻¹ yr⁻¹) to present acidification of podzolic soils (1-2 kmol hectare⁻¹ yr⁻¹).

Total proton sources and sinks refer to summed net H⁺ sources and sinks for various categories of processes. It is these net H⁺ fluxes, and not gross H⁺ transformations, which define the sign and magnitude of the irreversible H⁺ flux constituting soil acidification or soil alkalinization.

In conclusion, H⁺ budgets account for essentially all processes relevant to soil acidification. Concern about acidification by organic matter^{8,9} probably originates from the concepts of soil acidity and lime requirement. These concepts are important when raising the pH of an acidic, highly organic soil for crop production, but are of little significance when comparing internal and external proton sources in ecosystems.

N. VAN BREEMEN*
C. T. DRISCOLL†
J. MULDER*

* Department of Soil Science
and Geology,
Agricultural University,
6700 AA Wageningen,
The Netherlands

† Department of Civil Engineering,
Syracuse University, Syracuse,
New York 13210, USA

1. Driscoll, C. T. & Likens, G. E. *Tellus* **34**, 283-292 (1982).
2. van Breemen, N., Mulder, J. & Driscoll, C. T. *Pl. Soil* **75**, 283-308 (1983).
3. van Breemen, N., Driscoll, C. T. & Mulder, J. *Nature* **307**, 599-604 (1984).
4. Likens, G. E., Bormann, F. H., Pierce, R. S., Eaton, J. S. & Johnson, N. M. *Biogeochemistry of a Forested Ecosystem* (Springer, New York, 1977).
5. Driscoll, C. T., van Breemen, N. & Mulder, J. *J. Soil Sci. Soc. Am.* (in the press).
6. Martin, A. E. & Reeve, R. J. *Soil Sci.* **9**, 89-100 (1958).
7. Hargrove, W. L. & Thomas, G. W. *Soil Sci.* **134**, 216-225 (1982).
8. Krug, E. C. & Frink, C. R. *Science* **221**, 520-525 (1983).
9. Rosenqvist, I. T. *Sci. Total Envir.* **10**, 271-272 (1978).

Phonetic analysis capacity and learning to read

BRADLEY and Bryant's¹ result showing that training children on sound classification and letter-sound correspondences improves their later reading performance is clearly an important one. It brings definite support to methods of reading instruction that disclose explicitly the principle of the alphabetic code. The authors might have stressed that aspect of their findings, rather than the effect obtained with the group trained solely on sound classification, which fell short of significance.

The phrasing of the authors' conclusion that the connection of reading acquisition to phonological awareness is a causal one is unfortunate, however. If 'causal' means a link that runs one way only, the implication is to deny the possibility of the inverse type of relation. There is evidence, in fact, that learning to read improves understanding of the phonological structure of speech and hence ability to manipulate it at the level of phonetic segments. We have found, for example, that first grade (6-year-old) children taught by a look-say method performed more poorly on the task of interverting initial and final phonemes of an utterance than children taught by a phonic method². Groups of first grade children equated for age at testing time

were found to perform better on phoneme deletion and addition tasks when tested 5 months rather than 2 months after beginning learning to read³. In another study, carried out in an underdeveloped area of a Southern European country, we have shown that subjects who had learned to read at adult age performed considerably better than illiterates of comparable social background on similar tasks⁴. This result has since been confirmed and extended for two fresh groups of subjects. Learning to read is, however, not the only answer to phonetic competence, and we have found recently that a training procedure similar to that of Bradley and Bryant¹ can accelerate acquisition of phonetic analysis in 5-yr-old pre-readers⁵. In our opinion, a more satisfactory formulation is that the link between the ability to analyse speech into phonetic segments and reading acquisition is a part-whole one; acquiring the complete skill implies that the component capacities have been acquired and acquiring independently some of the component capacities can facilitate acquisition of the whole skill.

P. BERTELSON
J. MORAIS
J. ALEGRIA
A. CONTENT

*Laboratoire de Psychologie Expérimentale,
Université libre de Bruxelles,
1050 Bruxelles, Belgium*

1. Bradley, L. & Bryant, P. E. *Nature* **301**, 419-421 (1983).
2. Alegria, J., Pignot, E. & Morais, J. *Mem. Cogn.* **10**, 451-456 (1982).
3. Alegria, J. & Morais, J. *Archs Psychol., Genève* **47**, 251-270 (1979).
4. Morais, J., Cary, L., Alegria, J. & Bertelson, P. *Cognition* **7**, 323-331 (1979).
5. Content, A., Morais, J., Alegria, J. & Bertelson, P. *Cah. Psychol. Cognitive* **2**, 259-269 (1982).

BRYANT AND BRADLEY REPLY—In our paper we established that 4- and 5-yr-old children detect rhyme and alliteration before they begin to learn to read. This shows that some phonological awareness precedes reading, which rules out the possibility that all such awareness is the result of learning to read. The fact that our measures of this early awareness were related to later success in reading and that training in rhyme and alliteration improved reading appears to show that the skill has an effect on learning to read. Bertelson and colleagues suggest that we have not excluded the possibility that the cause-effect relationships go the opposite way and that the experience of learning to read improves children's awareness of the constituent sounds in words. We agree that we have not excluded this possibility—we did not intend to or claim to, and we think it likely that the causal links operate in this direction also.

There are some tests of phonological awareness which have proved more difficult than ours. One, used by Bertelson and his colleagues, is the elision task where a child has to work out what a particular word would sound like if a

phoneme in it were removed. Children do not manage this task at all well until some time after they have begun to read¹. The fact that there is such a large gap in the ages when children can manage our task and theirs suggests that the two tasks test different levels of phonological awareness, and that one level may precede and cause reading while the other may follow and be the result of learning to read. The possibility of different levels of the skill in question has not been raised before. We entertain the hypothesis that the basic difference between our easy test and the difficult one of Bertelson is that in ours the child has only to be aware of part of a word's sounds—the part that rhymes or that begins the word—while their test involves awareness and manipulation of all the word's sounds. We intend to test this hypothesis.

We thank Bertelson, Morais, Alegria and Content for several very lively and cordial conversations which ended in a substantial measure of agreement between them and us. These exchanges helped us to formulate our idea that some forms of phonological awareness precede reading while others follow it.

P. E. BRYANT
L. BRADLEY

*Department of Experimental Psychology,
South Parks Road,
Oxford OX1 3UD, UK*

1. Bruce, D. J. *Br. J. educ. Psychol.* **34**, 158-170.

Archaeomagnetism, Santorini volcanic eruptions and fired destruction levels on Crete

In their study of the magnetic properties of the Minoan pyroclastic deposits of Santorini and the fired destruction levels on Crete, Downey and Tarling¹ conclude that there was a time gap of several years or longer between the Plinian phase and the later phases of the eruption. They further propose that the Plinian eruption was contemporaneous with the destruction levels in central Crete, whereas the overlying base surge and pyroclastic flow deposits were contemporaneous with the destruction levels in eastern Crete. The geological and volcanological evidence, however, supports a single eruption with no time gap. I now suggest that their archaeomagnetic data should be interpreted in other ways more consistent with the other evidence.

The Minoan pyroclastic deposits are typical of those produced by many major explosive eruptions associated with caldera formation. An initial Plinian air-fall pumice deposit mantles the island and is covered by a complex sequence of base surge, pyroclastic flow and reworked volcanoclastic deposits². Similar sequences are known throughout the world^{3,4}. Large historical eruptions comparable in magni-

tude, style and geology include Vesuvius (AD 79), Katmai (1912), Tambora (1815) and Krakatoa (1883). All these eruptions produced a sequence of pyroclastic deposits, comparable with Santorini in periods of a few hours to a few days with Plinian and subsequent flow deposits often generated on the same day. Evidence indicates that large explosive eruptions associated with caldera formation are very short in duration and are not characterized by substantial time gaps.

Geological observations on Santorini² provide convincing evidence of a single short eruption. The Plinian pumice deposit forms a continuous sheet with a smooth upper surface across all the islands of Santorini. A prominent ash band (5-10 cm thick) occurs ~20-40 cm below the top surface of the Plinian deposit and can be traced around the entire caldera wall. The ash band thickens towards the north where, in the Oia region, it can be traced into a base surge horizon up to 1 m thick. This ash band demonstrates that base surge activity began before the Plinian phase was completed. There is no evidence of any erosion of the upper surface of the Plinian deposit and gullies or other disconformities are not observed. Major gullies, several metres thick, were cut through the Mt St Helens deposits within weeks of the 18 May 1982 eruption. There is no evidence of any features attributable to the rapid erosion of unconsolidated pyroclastic materials within the major part of the Minoan sequence. Channellized sheets of fluvial deposits are observed within the uppermost parts of the ignimbrites², but these are stratigraphically well above the horizons where the differences in magnetic properties are recorded. The Plinian deposit preserved its uneroded surface, because it was immediately buried by base surge deposits. A gap of many years cannot be reconciled with this evidence.

A major time gap should be obvious in the stratigraphy of the ash layer preserved in deep-sea cores⁵. In distal areas, east of Crete, the ash layer occurs as a single normally graded deposit. Two events, separated by many years or more, should produce two normally graded layers. In more proximal regions, the deep-sea ash layer is more complicated due to slumping into the Cretan basin⁵, but is also inconsistent with the time gap hypothesis. For example, the turbiditic layers are mixtures of ash from all the eruption phases, showing that slumping occurred after the ash layers had been deposited.

The geological evidence and volcanological arguments suggest that an alternative explanation of the magnetic data is required. One possibility arises from the emplacement temperatures deduced from the magnetic data. Downey and Tarling estimate temperatures in the range 300-400 °C for the base surge and massive upper deposits in Thera quarry. The deposits at Thera quarry are 40 m thick and would, therefore, have taken a con-

siderable time to cool. The cooling time of a sheet can be estimated from the calculations of Jaeger⁶ on the conductive cooling of igneous rock bodies. The outer margins of such a sheet, including the basal Plinian deposit, would cool quickly and would acquire their magnetic properties over a short time period at an early stage. The interior of the sheet will be insulated and take much longer to cool, and acquire the magnetic properties.

The time scale, t , for conductive cooling is related to the thermal diffusivity, k , and the sheet thickness, $2a$, through a non-dimensional parameter, τ by $\tau = kt/a^2$. Jaeger⁶ gives profiles of temperature for different values of τ through a sheet. For $a = 20$ m and $k = 8 \times 10^{-7} \text{ m}^2 \text{ s}^{-1}$, the centre of the sheet would remain at the initial emplacement temperature ($\sim 400^\circ \text{C}$) for just over a year. The centre of such a sheet would take 80 yr to cool down to 125°C and the time over which its magnetic properties were acquired would be much greater than the basal Plinian deposit lying directly on top of cold country rock. Cooling rates would be greater if convective processes (for example, rainfall) affected the porous pyroclastic sheet, but different parts of the deposit would still acquire their magnetic properties at substantially different cooling rates. The differences in magnetism could thus be explained by a prolonged and complicated cooling history rather than by a time gap.

R. S. J. SPARKS

Department of Earth Sciences,
University of Cambridge,
Cambridge CB2 3EQ, UK

1. Downey, W. S. & Tarling, D. H. *Nature* **309**, 519–523 (1984).
2. Bond, A. & Sparks, R. S. J. *J. geol. Soc. Lond.* **132**, 1–16 (1976).
3. Sparks, R. S. J. & Wilson, L. J. *J. geol. Soc. Lond.* **132**, 441–452 (1976).
4. Sparks, R. S. J., Self, S. & Walker, G. P. L. *Geology* **1**, 115–118 (1973).
5. Sparks, R. S. J., Brazier, S., Huang, T. C. & Muerdter, D. *Mar. Geol.* **54**, 131–167 (1984).
6. Jaeger, J. C. in *The Poldervaart Treatise* Vol. 2 (eds Hess, H. H. & Poldervaart, A.) 503–536 (Wiley, New York, 1968).

DOWNEY AND TARLING REPLY—Our results are an example of scientific serendipity, that is, the collection was undertaken assuming the volcanological consensus of a single eruption and the archaeological views of a single late Minoan IB (LMIB) destruction. We originally accepted Bond and Sparks' conclusion¹ that most of the tephra were of mud-flow origin, now known to be an ash flow, and hence W.S.D. sampled the tephra more than was strictly necessary. Our object had been simply to test whether these two catastrophes were synchronous and hence causally related. The evidence for a 10–20 yr gap necessarily came after the fieldwork and we have no finance to reinvestigate the field evidence. Nonetheless, we are surprised that Sparks should exclude all possibility of a gap. The main evidence cited is (1) the absence of erosion between the two, and (2) a Plinian ash

Table 1 Archaeointensity data		
Site and sample	Refs 2, 4	Ref. 1
H. Triada kiln, central Crete	$60.6 \pm 7.3^*$	68.1 ± 2 , $n = 6^\dagger$
Phaistos 1 kiln, central Crete	$\begin{cases} 37.8 \pm 6.4 \\ 89.6 \pm 15 \end{cases}$	67.3 ± 6.6 , $n = 4$
K. Zakro 1 kiln, eastern Crete	61.7 ± 3.8 , $n = 3$ [1, 2]‡	
K. Zakro 3 tile brick, eastern Crete	$\begin{cases} 57.8 \pm 2.5$, $n = 2$ [1, 2, 3] \\ 60.2 ± 2.5 , $n = 1$ [1, 2] (by Thellier technique) \\ 55.3 ± 0.8 , $n = 1$ [1, 2] (by ARM technique) \end{cases}	60.7 ± 2.4 , $n = 6$ (61.9 ± 5.7 for Makrygialos)
K. Zakro 1+3	$\begin{cases} 60.1 \pm 3.9$, $n = 5$ [1, 2] \\ 57.1 ± 5.2 , $n = 12$ [1, 2] \end{cases}	

* Standard error of the averaged results.

† n , Number of samples.

‡ Numbers in brackets indicate reliability: 1, very good; 2, moderately good; 3, dubious, combining Thellier and ARM (anhysteretic remanent magnetization) techniques as stated.

band that is traced into a base surge deposit. We agree that paroxysmal eruptions are of short duration, probably a few hours or days, but this does not exclude a time gap between the Plinian and later explosive eruptions. However, the short duration, combined with the magnitude of the eruption, implies that the paroxysmal blast was of high velocity so would readily move the light, unconsolidated Plinian pellets, leaving a 'sand-blasted' surface which could be difficult to distinguish from a pristine surface, particularly as the base surge probably eroded, as well as redeposited, earlier pumice in a way similar to that of Mt St Helens² (some 20 times smaller magnitude), so local, redeposited Plinian ash should be expected. We agree that the distal deep-sea 'Minoan' tephra is probably Plinian³. Nearer to the source, paroxysmal ashes must overlie Plinian, but seismic activity probably disrupted local layering and enhanced slumping at greater distances. Sparks' thermal model predicts that the magnetization would be 'frozen' in at different depths at different times. We do not find this. The directions are identical throughout the upper 60 m of upper tephra, which also has emplacement temperatures of some 400°C near the caldera, decreasing to 300°C on the flanks. Only a superficial few millimetres of large lava ejecta have been heated by contact with the tephra, suggesting rapid cooling—presumably by penecontemporaneous rainfall. We consider that our findings are consistent with the available evidence although further study and evaluation are desirable. In particular, analyses of the quantities of LMIB pottery at Cretan sites may be particularly informative.

W. S. DOWNEY*
D. H. TARLING

1. Bond, A. & Sparks, R. S. J. *J. geol. Soc., Lond.* **132**, 1–16 (1976).
2. Hoblitt, R. P., Miller, C. D. & Vallance, J. W. *U.S. geol. Surv. Prof. Pap.* **1250**, 401–419 (1981).
3. Pichler, H. & Friedrich, W. L. *Thera and the Aegean World* Vol. 2, 15–30 (Thera and the Aegean World, London, 1980).

DOWNEY AND TARLING assert¹ that the Plinian ash found in late Minoan (LM) Akrotiri is contemporary with the LMIA destruction on central Crete and that the ashes overlying these deposits in Akrotiri are, in turn, contemporaneous with the LMIB destruction in eastern Crete.

It is interesting to see the archaeomagnetic method applied to mud bricks and that such good results could be obtained. However, there is insufficient evidence, and the determination of a time gap between the above events has been over-optimistically stated by Downey and Tarling.

It is worth comparing the data in Downey and Tarling's Table 1 with other intensity and directional measurements which have been made^{2–4} on fired kiln clays from Phaistos (Ph), Hagia Triada (HT) and Kato Zakros (KZ) (see our Table 1). The apparent difference in magnetic intensity between HT and KZ is emphasized by Downey and Tarling more than it ought to have been, when the associated errors and techniques used in their case and in refs 2 and 4 are considered (incorporating the data in their Table 2).

Regarding the intensity variations of single measurements, the upper and lower error limits of 50 – $68 \mu\text{T}$ for eastern Crete (55 – 63 for ref. 2) and 55 – $73 \mu\text{T}$ for central Crete (53 – 68 for ref. 2) compared with 57 – $67 \mu\text{T}$ for Plinian ashfall, do not allow us to assess positively the two eruptive phases of the volcano.

The archaeodirectional data^{1,4} are compared in Table 2. Although barely discernible differences are noted in inclination (I°) data for HT and KZ from refs 1, 3 and 4 (which agree to within $\pm 1\sigma$ error), a difference in I° and D° (declination) observed for Phaistos, which is also noted in the intensity data (Table 1), may be explained by the use of measuring apparatus *in situ* as reported in Table 3.

The errors for the directional data of our Table 2 and Downey and Tarling's Table 2, and the degree of variation between them, indicate that this time

Table 2 Archaeodirectional data

Site	Refs	Declination, D°	Inclination, I°
H. Triada	3, 4	—	Mean $= 58.7 \pm 0.6$, $n = 2$
H. Triada	1	353.8 $\alpha_{95} = 2.3$	60.1, $n = 17$
Phaistos	3, 4	15.6 ± 3.3	56.8 ± 0.1 , $n = 4$
Phaistos	1	356.5 $\alpha_{95} = 1.7$	61.1, $n = 17$
K. Zakro	3, 4	—	56.6 ± 3.3 , $n = 5$
K. Zakro	1	355.8 $\alpha_{95} = 1.9$	54.6, $n = 49$
Makrygialos	1	354.3	57.5

Table 3 Differences between Sun and magnetic compass measurements

Site	Sun	Magnetic	Difference
Phaistos 3, 1	12.1	356	16.1*
Phaistos 3, 3	4, 8	348	16.8
K. Zakros 2, 1	256.7	194	62.7

* Compare this with Table 2, Phaistos D° values of refs 3, 4 and ref. 1.

difference is again indistinguishable. The directional differences for central and eastern Crete of $\Delta D = 1.7^\circ$, $\Delta I = 5.2^\circ$, with $\alpha_{95} = 3.5^\circ$ are not changes which support the hypothesis of Downey and Tarling. However, the *in situ* method for measuring inclination and declination should be borne in mind. Downey and Tarling report the appropriate errors for each sample which reduced after averaging to the order of $1-2^\circ$. Whenever possible they used a Sun compass or, failing that, a magnetic compass.

We have found that declination results from samples orientated using a magnetic compass should be considered, in certain cases, as extremely dubious^{3,4}, as they are affected by stray magnetic fields present in a magnetized structure such as a kiln, many pithoi, burnt mud-brick walls, metal fences or an immediate volcanic environment. (The inclination results are not affected by the errors in present magnetic orientations, but could have been affected if the original cooling were in a high magnetic environment.) A comparison of D° values using Sun and magnetic compasses respectively showed differences in the degrees (see Table 3). The Sun compass provides more accurate results.

Other possible effects that might alter the reliability of measurements of D° and I° include earthquake activity in the Aegean seismic arc, which produces tilting with possible local subsidence; thus the tectonic instability of the Knossos region in central Crete with respect to eastern Crete has been well documented⁵. Such land movements⁶ must also have an effect on the supposed stable burnt structures and ash materials.

Moreover, the repeated and pronounced volcanic (and seismic) activity of the island in ~ 197 BC and AD 1950 could have produced discernible shift effects in the

original tephra deposits. Furthermore, the possibility of a marked degree of reheating of the base surge and upper tephra deposits in Santorini through these repeated violent explosions⁷ should be examined.

Referring to the secular variation of the geomagnetic field for this period, other workers⁸⁻¹² suggest that around 1500-1100 BC, a change in the magnetic field of up to 50% of the present value for Crete (44 μ T for 1975) took place. As such, the present evidence for the rise of change in direction and intensity (for the same period), with their associated errors of $\pm 2^\circ-5^\circ$ and $\pm 5\%$ respectively, would give a dating accuracy comparable with the time gap in question.

Another attempt to monitor successive eruptions has been made by applying U-Th series radioisotopic analysis to tephra deposits from Kos, Rhodes, Nisyros, Tilos and Santorini¹³. Differences in the radioisotopes content of the tephra were found which reflect a different magma source. Results indicated that a second eruption could not be ruled out, although not one that was necessarily derived from Santorini itself.

I thank Professor K. Creer, Dr M. Aitken and G. Bussell for reading the manuscript, and the National Hellenic Research Foundation-Royal Society of London Exchange Scholarship.

YIANNIS LIRITZIS

Ministry of Culture and Science,
Department of Underwater Antiquities,
58 Omirou St,
Athens, Greece

DOWNEY AND TARLING REPLY—We agree with most of the comments of Liritzis. The palaeointensity observations are subject to a range of errors, therefore we quoted these observations as supporting the directional evidence, rather than using them as primary indicators of age differences. We also concur with his comments concerning the use of a Sun compass, but it is not possible to use it at several key sites, such as Akrotiri, where the structures are protected by a solid roof. In other areas, the Greek authorities were extremely cooperative in removing covers wherever possible. Comparison of Sun and magnetic compass readings on open sites of fired mud bricks showed no differences between the readings, but we are also suspicious of sites with magnetic orientation alone, particularly where large structures are concerned. Luckily, the main difference in direction is in inclination and this tends to be less affected by errors in determining the direction of strike lines. Partially because of such problems in kilns, few of our results were included from these structures (the main exception being the only site we were allowed to sample at Knossos), although the main reason for this omission was that

the normal firing of these may not be synchronous with the natural(?) destructions of houses, palaces, villas, and so on. We also accept the possibility of regional tilting, although the coastal evidence for such tilting has a very low angular value. In terms of the volcanics, both tephra sequences would be tilted if significant seismic movements had occurred, while it is difficult to see why subsequent thermal effects should be so confined to specific layers and be so uniform within them that no other components can be isolated. We emphasize that any interpretation would benefit from more intensive and extensive study, including presently unsampled localities, but we were convinced by the difference in direction, supported by palaeointensity differences, which emerged from the results during the final stages of analysis and, on reviewing the problem, the short time gap seemed to offer a possible explanation of other perplexing features, such as differences in the quantities of LMIB pottery retrieved from different sites.

W. S. DOWNEY*

D. H. TARLING

Department of Geophysics
and Planetary Physics,
University of Newcastle-upon-Tyne,
Newcastle-upon-Tyne NE1 7RU, UK

* Present address: Department of Physics, University of Papua New Guinea, Port Moresby, Papua New Guinea.

- Downey, W. S. & Tarling, D. H. *Nature* **309**, 519-523 (1984).
- Liritzis, Y. & Thomas, R. *Nature* **283**, 54-55 (1980).
- Liritzis, Y., McKerrell, H. & Creer, K. M. *Archaeometric Sampling in Greece* (Field Rep., Dept of Physics and Geophysics, Edinburgh University, 1977).
- Thomas, R. thesis, Edinburgh Univ. (1981).
- Pichler, H. & Schiering, W. *Nature* **267**, 819-822 (1977).
- Liritzis, Y., Miserlis, P. & Rigopoulos, R. *Int. J. naut. Arch.* **12**, 191-202 (1983).
- Rittman, A. *Volcanoes and their Activity*, 3-5 (Wiley, New York, 1962).
- Games, K. P. *Geophys. J. R. astr. Soc.* **63**, 45-56 (1980).
- Walton, D. *Nature* **277**, 643-644 (1979); **310**, 740-743 (1984).
- Aitken, M. J., Alcock, P. A., Bussell, G. D. & Shaw, C. J. in *Geomagnetism of Baked Clays and Recent Sediments* (eds Creer, K. M., Tucholka, P. & Barton, C. E.) 122-127 (Elsevier, Amsterdam, 1983).
- Kovacheva, M. *Geophys. J. R. astr. Soc.* **61**, 57-64 (1980).
- Shaw, J. *Geophys. J. R. astr. Soc.* **58**, 107-116 (1979).
- Liritzis, Y., Ivanovich, M. & Gale, N. H. *Int. Conf. Archaeometry*, Naples (Leric Foundation, Rome, 1983).

Matters Arising

Matters Arising is meant as a vehicle for comment and discussion about papers that appear in *Nature*. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted. Neither contribution nor reply (if one is necessary) should be longer than 500 words and the briefest of replies, to the effect that a point is taken, should be considered.

Memories of the future

Igor Aleksander

The Creative Computer: Machine Intelligence and Human Knowledge.

By Donald Michie and Rory Johnston.

Viking: 1984. Pp. 263. £12.95, \$29.95.

The Cognitive Computer: On Language, Learning, and Artificial Intelligence.

By Roger C. Schank and Peter G. Childers.

Addison-Wesley: 1984. Pp. 268. \$17.95.

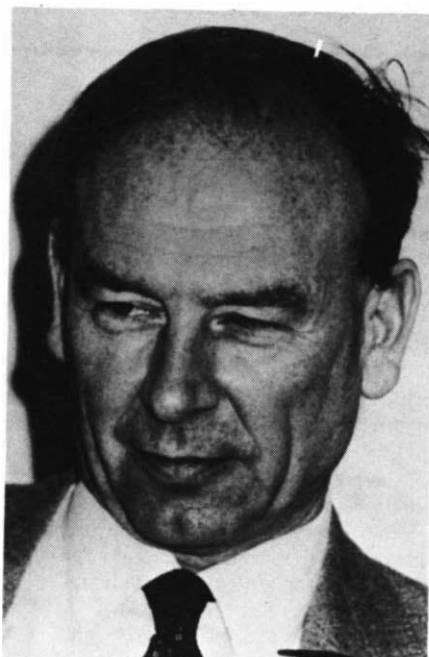
ALTHOUGH artificial intelligence has already completed its first quarter of a century as a recognizable discipline, it is only now on the brink of capturing the attention of a wider audience. Despite the age of the subject it still carries with it a set of contradictory and confusing notions. The arguments are not about the technicalities (which, all would agree, contain some pretty clever programming tricks), but about *why* anyone should be interested in AI at all. One of the authors of each of these two books is an innovator of AI ideas. Donald Michie has, for many years, led the AI movement in Britain and has contributed internationally to the design of "heuristic" programs that rapidly find their way through "trees" of complex decisions. Roger Schank is one of the founders of the Yale AI school, where some of the fundamental programs were written that enable machines to react appropriately to human-like conversations.

In the two books, neither author refers to the work of the other indicating that the reader is offered two very different views of what the subject of AI is and where it is going. Michie and Johnston suggest that AI may spawn machines which, being smarter than human beings, will provide ways for mankind to pull back from the brink of disaster. Schank and Childers, on the other hand, argue that the central aim of AI is to provide systems that understand and obey their human users without making excessive demands on the way that a person wishes to express himself to the machine. Their message is one of banishing the pains of learning to work with the patently unintelligent but ubiquitous programmed computer of today. In other words, Michie is happy to suggest that machine intelligence is not only like the human version but also an improvement of it — while Schank sees the two as being very different, but cooperative.

Michie and Johnston's case rests on the proposition that a machine can be *creative*, in the sense that it can get to "know" things that were not directly fed into it by a human programmer. They go to some trouble to explain that "expert system" programs are dragging the computer out of the dark ages into a dawn of supreme enlightenment. An expert system program includes a long list of logical rules and facts about some field of expertise (for example, oil prospecting). It is then capable of being interrogated by a

non-expert and of suggesting answers (for example, whether an area of land is worth exploring or not) after a question-and-answer interaction. The system not only shuffles around its data to provide the answers, but also explains how it got there

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS



Roger Schank (top) and Donald Michie — different views of what AI is and where it is going.

by backtracking through the logical rules it has used.

Michie's case is that as a result of "Fifth Generation" research such systems will contain prodigious amounts of memory. This will hold the expertise of many experts in many fields and become super-smart through the sheer volume of stored knowledge. Readers will soon realize that this is the same as arguing that if all the encyclopaedias in the world could be stored in a machine and accessed rapidly, a super-wise machine would result: a knowledgeable machine, yes; but one that was better able to add wisdom to, say, the nuclear disarmament debate, no. That kind of creativity is not contained in encyclopaedic knowledge: it requires a leap of insight outside itself which, as yet, no expert system has achieved.

But Johnston and Michie are aware of this and try to patch up the gap by producing, as evidence, programs that have found shortcuts in planning problems not fed in by the rule-supplying human expert (these are commonly found in programs which play games). In less than a page of the entire book they point to a program called ID3 (developed within Michie's team) which generates advice tables about the winning potential of certain chess end-games by correlating a wide variety of board patterns with possible outcomes. This is creativity of the kind that is found in the interference-reduction engineering systems carried on board exploration satellites. The signals received on Earth from the satellite's cameras suffer from distortion and disruption due to random "static". But computers using statistical techniques remove this static, revealing beautiful, clear pictures never before seen by man! In a similar way, the advice tables have revealed order not previously seen by human chess players. But the case that such laundering operations will lead to "wiser-than-man" computers requires an act of faith to which only the most enthusiastic readers may subscribe.

In contrast, Schank and Childers's approach is refreshingly realistic and credible. The central question they see as being whether sufficient "knowledge" may be included in a computer so that its responses may be helpful, while reflecting the fact that the machine has no "opinion" of its own. Such machines might be expected to absorb stories in natural language (reports from the scene of an accident, for example), paraphrase them into essential facts and answer questions. In a story that contains something like "Jack ate a bacon sandwich which he enjoyed and washed down with a pint of beer", it would be difficult but possible for the machine to answer correctly a question about whether eating or drinking took place first. In answer to a question such as "Are bacon sandwiches tasty?", the appropriate machine response might be

"Jack must have found the sandwich in the story tasty, but I cannot answer the question in general as I do not eat bacon sandwiches".

The book contains a progression of convincing examples that point the way towards such interactions, while continuously reminding the reader of the enormous programming difficulties that this implies. For Schank the massive memories of the future will be used to make helpful, interactive systems from which people could learn world knowledge in fields that are new to them. He does not believe in the super world-saver but sees systems as developing into patient and understanding teachers. The latter parts of *The Cognitive Computer* return to the positive benefits of this sort of AI: better health services, better banking systems and even better democracies.

There are two economic and social reasons which suggest that Schank's view is right. The first, agreed by both authors, is the falling costs of knowledge storage. The second is the powerful market edge of a machine which is well-tuned to human use over one that isn't. In other words, no industrial concern could bear the costs of Michie's megalomaniac machine and hope to recoup the outlay from a free-economy market. Schank's helpful cooperator, on the other hand, sounds far more marketable than even the most successful micro in the High-street shops today.

On a critical note, it is a sad fact that neither of these books says much about the unprecedented opportunities which lie ahead for new computer structures and designs encouraged by the silicon chip revolution. Both sets of authors pin their hopes on machines whose general structure has hardly altered since 1947. For example, several laboratories in the world are exploiting low chip costs by researching parallelism (a trick that the human brain uses to advantage). This may not only be the precise vehicle that brings down computing times to fulfil some of the predictions, but also opens new opportunities for a flexible and rapid acquisition of knowledge which on conventional machines would just take too long.

If readers can stand up to the continuous barrage of suggestions by Michie and Johnston that future survival depends on machines which may never be built, they may well find the historical approach and the lively style of *The Creative Computer* acceptable. In contrast, Schank and Childers's *The Cognitive Computer* is written in an appealing, absorbing and highly informative style throughout. With its clear philosophy and lucid examples of some difficult programming concepts, it ranks as one of the best general books on AI written to date. □

Igor Aleksander is Professor of the Management of Information Technology in the Department of Computing, Imperial College, University of London.

Variable climates

B.G. Hunt

The Global Climate.

Edited by John T. Houghton.
Cambridge University Press: 1984.
Pp.233. £27.50, \$49.50.

ON THE tails of the surge of interest in climate has come a number of books dealing with various aspects of the subject. *The Global Climate* is one of the most recent, and tries to encapsulate current knowledge by drawing on the talents of a number of scientists working in constituent areas of climatic research. The preface states that it is designed "to explain to a wider scientific community the background, aims and main lines of research being pursued". It is to be hoped that this is not the only justification for the book, because, apart from the general summary in the first chapter, it will probably appeal only to those already active in the subject.

The 12 contributions are somewhat specialized in nature and range over climatic observation, modelling, clouds and radiation, ice, the oceans and atmospheric composition. There is no index – a severe drawback – and the standard and content of the chapters are highly variable, some being overly detailed and others quite short and specific. For example the observational study is essentially confined to discussing variations in temperature, which hardly illustrates the real complexity of the observed climate. By contrast the elaborate description of model experiments concerned with soil moisture and surface albedo variations includes virtually every such experiment performed; this detail is hardly justified when no real ap-

preciation is given of the very poor representation of soil moisture used in models at present. There is a similar problem with the account of clouds and radiation which is mainly about results from one-dimensional models, far too little space being devoted to the real problems of modelling the three-dimensional world.

The three chapters on oceanography highlight the important role that the oceans play in climate, and emphasize the complexities of the interactions between the oceans and the atmosphere. Wood's contribution on the upper oceans is particularly comprehensive, but is marred by his excessive use of references; at times the article reads like a telephone directory. The following chapter by Wunsch, on the ocean circulation, is the most stimulating in the book. He emphasizes the problems which have to be resolved, and the book as a whole would have been all the better had more of the authors taken this tack.

The Global Climate portrays, warts and all, the very great complexity of the climatic system and the tremendous problems involved in trying to predict the future. It would have been improved by chapters on palaeoclimatology, and particularly on theory, where the underlying dynamical principles of the climatic system could have been illustrated. But what, above all, is generally lacking is any sense of the excitement offered by climatic research. Communication of this point to a wider audience is necessary if more first-class brains are to be attracted into the field. The intellectual challenges are certainly there – prospective high-energy physicists please note. □

B. G. Hunt is a Senior Principal Research Scientist at the CSIRO Division of Atmospheric Research, Melbourne.

All about polymers

M.V. Volkenstein

Macromolecules, 2nd Edn. Vol. 1 Structure and Properties; Vol. 2 Synthesis, Materials, and Technology.
By Hans-Georg Elias.

Plenum: 1984. Vol. 1 pp.521, \$65, £50.25; Vol. 2 pp.821, \$95, £80.75.

MACROMOLECULES are the building blocks of life, and study of them unites the three prime natural sciences – chemistry, physics and biology. On the other hand, macromolecules in the form of polymers clothe not only people but the wheels of cars and planes, and are used as substitutes for metals and glasses in industry. The writing of a comprehensive, up-to-date account of macromolecular science and technology is thus a complicated and difficult job, but that goal has been achieved by Dr Elias. The original edition of *Macromolecules* acquired a well-

deserved reputation for thoroughness and reliability; this second edition, written practically *de novo*, is much more complete and should be even better received.

The two volumes are encyclopaedic in scope and contain an enormous amount of information, beginning with the chemical structure of macromolecules, their configurations and conformations, and ending with the description of specific materials such as fibres and adhesives. The reader will find here information on nearly everything to do with polymers, from a simplified derivation of the formula for the dimensions of the polymer coil (p.119) to data on the annual production of oil in different countries (p.857). (Oil is one of the best sources of raw materials for polymer production; Mendeleev wrote that the burning of oil was the same as the burning of currency bills!) The whole is written in a clear and straightforward style. And while the approach reflects Dr Elias's viewpoint, the reference lists include all the most important papers, reviews and monographs on the subject.

Dr Elias himself specializes in macromolecular chemistry — the mechanisms and kinetics of various types of polycondensation and polymerization — so one would expect the chemical aspects to be described especially carefully, as indeed they are. The same can be said about the technological side, the characteristics of all the classes of polymers known to date being given.

The coverage of the physical aspects, however, while good is not quite so thorough. Two decades ago many of us believed that physics of polymers had reached a dead-end and that there were only two ways forward, either towards biology or towards technology. Later it became clear that we were wrong. New ideas were proposed in the theory of macromolecules based on the analogy between the properties of polymer systems and systems fluctuating near the critical points of phase transitions. These ideas were put forward on one hand by P.-G. De Gennes, and on the other by the late I.M. Lifshitz and his school.* One of the most important places in this new physics belongs to the theory of the coil-globule transitions. It is a pity that Elias does not tell us anything about this field, and does not describe the specific physics of the globular state. Equally he does not give an adequate treatment of the theory of the relaxation properties of polymers, which forms the background to the understanding of their electrical and mechanical properties; nor, on the topic of methods, does he mention the circular dichroism of chiral polymers which replaced the study of their optical activity long ago.

From my personal point of view, the description of biopolymers and related biological problems also falls short of what would, at best, be desirable — graduate students of polymer chemistry should know more about this field than is discussed here. There are also a few mistakes; for instance triple helices can be formed by synthetic polynucleotides but not by natural nucleic acids.

But of course the work of one author, especially when it runs to 1,342 pages, cannot be totally free from some second-hand defects; and we cannot criticize a pear tree for producing pears but not apples, though these points might perhaps be taken into account for the next edition. As it stands the fruits of this massive work are very rich indeed. *Macromolecules* will find a wide circle of readers among graduate and postgraduate students, scientists and engineers. □

M.V. Volkenstein is at the Institute of Molecular Biology, Academy of Sciences of the USSR, Moscow.

*See, for example, P.-G. De Gennes *Scaling Concepts in Polymer Physics* (Cornell University Press, 1980) and I.M. Lifshitz, A.Y. Grossberg and A.R. Khokhlov *Rev. Mod. Phys.* 50 N3, 683 (1978).

Publishing history (and more)

T.G. Rosenthal

Cambridge University Press 1584–1984.

By M.H. Black.

Cambridge University Press: 1984.

Pp. 343. £12.50, \$19.95.

AS A commercial publisher, educated almost exclusively in the humanities and subject to advanced bibliomania, for me to review this book is rather like a satyriasis sufferer being asked to pronounce on the bound volumes of *Playboy*. I must also declare a two-fold interest in that I was an undergraduate at Pembroke College, Cambridge, under the Mastership of S.C. Roberts — one of the most influential Secretaries Cambridge University Press has had — and the author of the book under review once offered me a job. But as that was nearly a quarter of a century ago, I can perhaps reasonably claim some objectivity today.

Most “house histories” amount to little more than hagiography and are read only by those with an *a priori* connection with the company in question. In theory, the story of four centuries of a publishing house should be of much wider appeal. Such is the case; the book is compelling reading for anyone with an interest in the spread and propagation of literature and learning as well as the history of publishing itself.

The threads that run through Black's account of the history of CUP are manifold and there is insufficient space here to do them justice. From the receipt of the royal charter in 1534, and the appearance of the first publication in 1584 (just one year ahead of Oxford), he tells of changing printing skills and technology down the centuries, of typography and typographers, of eccentrics and bores, of legal constraints on publishers and battles between them, and a great deal more, much of it of consequence far beyond Cambridge itself. Perhaps the most salient and continuous theme, however, is the development of bible printing, worth a book in itself and an activity which, throughout its history, has brought enormous profits to the Press thus enabling it to develop its traditional scholarly publishing.

If I have any serious quarrel with CUP it is here: in producing jointly with Oxford University Press the New English Bible, the Press produced at a cost of more than one-and-a-half million pounds a book utterly lacking in distinction. It is an unenviable mark of philistinism that two of the three licensed English bible publishers (the other is Eyre & Spottiswoode) should have combined in so massive an effort to provide so mediocre an alternative to the King James's Bible (which, even to a Jewish agnostic like myself, remains one of the masterpieces of

religious writing as well as a great work of literature in its own right).

That opinion is unlikely to be adhered to within the Press, at least not officially. However, in the book Black is unequivocal about an earlier decision, of 1877, that of the Syndicate of the day not to publish what later became the *Oxford English Dictionary*. “Possibly the largest wrong decision in publishing history” he calls it, but is understandably philosophical about the matter:

It would be a cautionary tale, if publishers did not know the moral from their own experience, and if it were possible to guard against being monumentally wrong by any simpler method than always being right.

However, the Press would not be alive today if it hadn't got most of its decisions right. By 1890 it was “recognisably becoming a modern publisher” and was commissioning some books rather than merely waiting for manuscripts to arrive. Beyond the bibles and the prayer books it had the makings of a humanities list, a scientific list and a schoolbook list. And, while many a publisher, put off by his appalling handwriting and idiosyncratic style, could well have rejected C.M. Doughty's *Arabia Deserta*, one of the more perceptive Syndics, the Arabist, William Robertson Smith, recommended publication subject to Doughty accepting Smith's editorial guidance “to render the work intelligible to ordinary English readers”. Poor Doughty was asked to contribute to the costs of publication, and those who have been dunned by their publishers for excessive tinkering with their proofs may take comfort from the fact that by publication day, in 1888, Doughty owed the Press £730 (surely not less than £15,000 in today's terms). Doughty was unwilling to pay and in the end surrendered his copyright instead; but lest anyone think the Press unduly rapacious, it was stipulated that the copyright could revert to him, after the first impression, upon payment of £200.

Towards the end of the nineteenth century, CUP, some 200 years previously the publisher of Isaac Newton, really hit its stride as a producer of scientific works. J.J. Thomson, C.S. Sherrington, Lord Kelvin, Lord Rayleigh, James Clerk Maxwell, Osborne Reynolds, George Stokes and Alfred North Whitehead all published under its imprint, while the first science journal put out by the Press, the *Journal of Physiology*, appeared in 1893. The stream continued through the turn of the century — with Rutherford's *Radio Activity*, Hardy's *Pure Mathematics* (published in 1908 and still in print) and the three volumes of Whitehead and Russell's *Principia Mathematica*, among many others — and on after the First World War. Alongside this galaxy of scientific works, other aspects of the Press's activities flourished. In particular there is the extraordinary record in the publication of history with what, to historians, are

household names — G.G. Coulton, Dom David Knowles, Sir Stephen Runciman, Sir John Clapham, Eileen Power and many more.

Altogether, the decades between the two World Wars may perhaps be seen as the greatest years in all of CUP's existence. Onto financial stability and imaginative publishing were grafted the typographical influence of Stanley Morison, making the Press a leader in Britain and the world in the combination of style with clarity in typographical design. There are many reasons for reading Black's book and among them is a series of brief but effective character sketches — of Morison, for example, quite obviously a most eccentric man despite (or because of) being a typographical genius; or of someone called Nobbs who, according to Black, was "curmudgeonly until excited, when he became virtually impossible".

Throughout its history the Press has had difficulties — financial and political — but none so severe as the crisis of the late 1960s and early 1970s which threatened to take it under. The extent of the problem is made clear by Lord Todd, appointed as Chairman of the Syndics in 1971, in his autobiography *A Time to Remember* (quoted by Black on p. 243):

Already on my first visit as Chairman...I was horrified: the Press was to all intents and purposes bankrupt, with a soaring overdraft and with sales and receipts dwindling.

Todd moved with remarkable speed and steered through the appointment of Geoffrey Cass who, as Chief Executive, must be credited with the reorganization that has made CUP as efficient and, I suspect, as profitable as it is today. (One can only suspect; while Black is by and large good on quoting figures, numbers of employees and so on, he is brief to the point of evasiveness on the current turnover and profit of the Press.)

Cass did many things that needed doing. Among them was the enforcement of the policy that current book prices had to reflect replacement costs, not original costs, a point it has taken far too many publishers far too long to realize during the past decades. Indeed, of the many constituencies in the potential audience of this book, the oddest, I suspect, must be the world of accountants. The master-stroke, however, came from a different direction. By brilliant research in the university archives and sophisticated presentation of his arguments, Cass managed to persuade the Inland Revenue that the Press should become a charity which would be exempt from payment of corporation tax. His 1975 submission to the District Inspector of Taxes consisted of some 60 typewritten pages, one of which contained the sentence "Our formal submission, when completed in the Spring, is likely to be several hundred times bigger...". Clearly Mr Cass is not a man to be trifled with. He succeeded in his aim when the Inland Revenue agreed on 9

Developments in holography

Brian J. Thompson

Optical Holography: Principles, Techniques and Applications.

By P. Hariharan.

Cambridge University Press: 1984. Pp. 319. £35, \$69.50.

IN HIS preface, Dr Hariharan states: "My aim in writing this book is to present a self-contained treatment of the principles, techniques and applications of optical holography, with particular emphasis on recent developments". He has succeeded admirably in that endeavour. It was time for a new, authoritative text on holography, to provide a succinct but readable introduction to the principles and current methodology of optical holography, and up-to-date coverage of applications.

The book contains 15 chapters and five short appendices. The first three chapters are concerned with the fundamentals, the first giving some historical perspective, the next two being accounts of the basic methods of hologram formation (in-line, off-axis, Fourier transform, Fraunhofer, and image) and of image formation from the hologram. The various types of holographic recordings (thin, volume transmission and volume reflection) are described in some detail in a separate chapter.

Three further chapters deal with many of the important practical aspects of the art and science of holography, including light sources and optical systems, parameters of recording media and practical recording media. Various useful chemical formulae and processing conditions are given for monobath developers, tanning bleach, reversal bleach, dichromated gelatin and so on.

Separate chapters are devoted to holograms for displays, colour holography, computer-generated holograms, information storage and processing and holographic interferometry. Special tech-

niques — including polarization recording, hologram copying and holography with incoherent light — are discussed as a group. (I confess I still have semantic difficulty with the phrase "holography with incoherent light", or worse yet "incoherent holography"!)

Imaging applications of holography are covered, among them microscopy, particle size analysis, imaging through moving scatterers and distorting media, correction of aberrated wavefronts, high-resolution projection imaging, evanescent-wave holography and multiple imaging. Rather oddly, this chapter also contains sections on holographic optical elements including scanners and diffraction gratings.

Some readers will be disappointed by the brevity of one or two of the sections; for example, the discussion of holographic gratings takes a bare two-and-a-half pages, including two figures. However, to compensate, there are ample references in all sections, some 700 in all, about 10% of which are from 1980–1982; the author has indulged himself a little with the largest number (32) to his own work — at least twice as many as the nearest competitor! There is also a short bibliography, an author index and a very good subject index, although I was surprised that multiplex hologram, zone plate, zone lens and far-field hologram are not included.

Topics such as speckle, interference and coherence are dealt with briefly in the appendices. Coherence requirements are important but unfortunately the author repeatedly writes about the spatial coherence of the source when he means the spatial coherence of the field produced by the source.

Such small defects do not affect the overall value of the book. It will appeal to a wide audience, from those embarking on a serious study of holography to the established researcher who will use it as a reference text. □

Brian J. Thompson is Professor of Optics in the College of Engineering and Applied Science and Provost at the University of Rochester, New York.

November 1976 that CUP was indeed tax exempt. The so-called older university, and younger university press over at Oxford, watched elegantly from the sidelines and subsequently procured the same exemption.

The extent of Cass's success and the present health of CUP can be judged from the current scale of the Press's activities. Nineteen eighty-four, it is estimated, will have seen from the publisher of Newton and Darwin, Dryden and Milton, some 900 new publications with projected sales of eight million units.

This is an altogether admirable book, hugely enjoyable and informative. It is superbly produced — the broken letter in the word "editorial" (p. 286) apart — and

so indeed it should be. Black's prose is lucid and frequently witty, while the work as a whole is far less self-serving than most volumes of a similar nature. My only real disappointment is the absence of the present-day figures that really matter. If companies who do pay their taxes have to tell all, should not charitable institutions also be compelled to report fully — if only to reassure those of us who toil elsewhere in the vineyards of publishing that the two great university presses are not thereby gaining a massive commercial advantage. □

T.G. Rosenthal is Joint Chairman and Joint Managing Director of André Deutsch Ltd, London.

New products described this month include three made with safety in mind—insulated gloves, a system for mopping up spilt chemicals and a wallchart on hazard warning symbols.

● LKB Produkter AB of Sweden has introduced a new embedding system for high-resolution light microscopy. Histo-Resin, which is a glycol methacrylate polymer, offers a rapid low-toxicity procedure for the production of highly transparent specimen blocks. The HistoResin kit includes all components required for specimen infiltration and embedding. Thin (2µm) serial sections can be cut for staining series, and the HistoResin system includes a special mounting medium to ensure stable blocks and secure attachment of virtually any adapter. Fully polymerized, clear blocks are ready for sectioning within 2 hours.

Circle No. 100 on Reader Service Card.



Chemops in action

● For fast and easy cleanup of chemical spills, Lab Safety supply Co. produces the Chemop. The mop works like an ordinary squeeze-action mop, but is made entirely of chemically resistant and nonsparking materials, and fitted with disposable Chemsponges that will absorb acids, caustics, solvents or oils. The Chemop eliminates the use of messy absorbents and neutralizers and greatly simplifies spill cleanup.

Circle No. 101 on Reader Service Card.

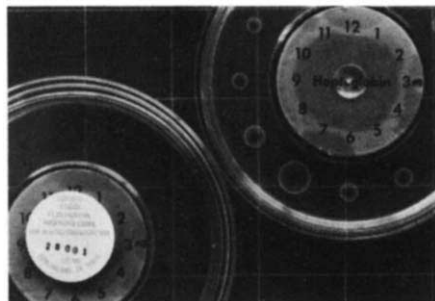
● Safetemp gloves, from Forma Scientific, protect the hands from the cold while working with cold or frozen materials. The heavy-duty water repellent gloves are effectively insulated yet the gloves are lightweight for good mobility. Safetemp gloves are available in three sizes — elbow, mid-arm and shoulder lengths.

Circle No. 102 on Reader Service Card.

These notes are based on information provided by the manufacturers. For further details circle the appropriate numbers on the Reader Service Card bound inside the journal.

● NOTEBOOK is a new software system from Biodata intended to do in the laboratory what the wealth of office computer software available does for the office. The new package provides a standardized and simple way of controlling experiments and collecting data. Data are communicated between the NOTEBOOK program through disk files which can also be transferred into files suitable for off-the-shelf office data processing packages. Programs are written in PASCAL for speed of execution and make extensive use of menus and soft keys to select program options. The programs will run on the IBM PC and compatible machines including Apricot and Sirius. Typical applications include long-term data logging — recording temperature, strain, analogue voltages; waveform capture and display; analogue-to-digital and digital-to-analogue conversion; off-line signal analysis.

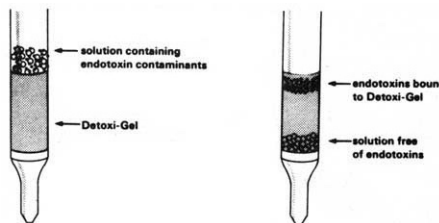
Circle No. 103 on Reader Service Card.



The Diffu-Gen in vitro diagnostic system

● The Diffu-Gen radial immunodiffusion assay system from Tago now includes an assay for measuring the concentration of haptoglobin. Tago's RID agarose plates are colour-coded, contain 12 numbered wells, and are available in 3- and 10-unit packages. A Reference Sera Kit of three levels of haptoglobin can be supplied separately. Tests can be performed without diluting patient sera, and no incubator or humidity chamber is required.

Circle No. 104 on Reader Service Card.



Detoxi-Gel removes endotoxins from solution

● Endotoxins can be removed from a solution using a new affinity chromatography support from Pierce Chemical. The support, Detoxi-Gel, uses a special ligand covalently bound to an agarose matrix and has a high capacity for endotoxins.

Circle No. 105 on Reader Service Card.



The fourth edition of the BDH Hazard Warning Symbols Chart shows the symbols used in current UK labelling regulations. The charts measure 70 × 49 cm. Circle no. 106 for further details.

● As a result of a recent cooperative agreement between ICI and the Institute of Radioelements, Fleurus, Belgium, scientists in the United Kingdom and the United States now have access to a Tritium Custom Synthesis Service by contacting the Physics and Radioisotope Services Group of ICI at Billingham. Compounds may be labelled by a variety of techniques including catalytic exchange, halogen replacement and reduction with tritium gas.

● Mung bean nuclease, recently reported capable of isolating genes by preferentially cleaving DNA at the ends of the gene segment (in the presence of formamide), is available from Pharmacia. Mung bean nuclease is also often preferred over S₂ nuclease for single-strand specific cleavage, as it has less tendency to nick double-stranded DNA.

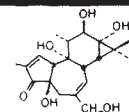
Circle No. 108 on Reader Service Card.

● A new 12-page booklet available from Data Corporation details the immunoperoxidase staining characteristics of a variety of normal and neoplastic species using ten monoclonal antibodies.

Circle No. 109 on Reader Service Card.

Advertisement

PHORBOL
ESTERS



LC SERVICES CORPORATION

165 New Boston Street
Woburn, MA 01801
(617) 938-1700

The importance of training

from Richard Pearson

A recent study suggests ways in which the United Kingdom might improve industrial competitiveness by investing in training.

THE United Kingdom's position in the world 'high-tech' league has become the focus for increasing concern over the past year. While the shortage of information technology skills has been the focus of much attention, and is a problem afflicting many other countries, a more disturbing manpower and skills problem for the United Kingdom has been highlighted in the latest audit of the country's performance, the report *Competence and Competition* (HMSO, 1984). This report seeks to contrast UK investment in training and human resources with that in West Germany, Japan, and the United States. Making international comparisons of educational qualifications and training investment can of course be notoriously difficult and lead to endless arguments, for example, over whether a Japanese engineering graduate is equivalent to one from West Germany or the United States. To get round this problem the report draws on a wide range of indicators to argue that the United Kingdom lags behind its major competitors in its investment in training and education, and it sets out an agenda for action to improve competitiveness and performance for the future.

While much of the recent national debate has highlighted the high tech skills, the report takes as the foundation of a successful vocational educational and training (VET) system, the preparation of young people. The high participation rates of young people in education in the three countries is highlighted. In Japan 94 per cent of young people stay in education until the age of 18, and 40 per cent go on to higher education or graduate school. In West Germany, nearly 70 per cent of school leavers go into VET and 20 per cent enter some form of tertiary education, while in the United States 73 per cent of those aged 16–24 are enrolled in school, college or university. By way of contrast, in the United Kingdom only just over half of the 16–18 age group participate in full-time education (Table 1). Emphasis on the 'three Rs'

and competences such as team work, flexibility and 'learning to learn' rather than occupational knowledge and skills are seen to be important ingredients of successful VET systems. In terms of specific qualifications, 87 per cent of Japanese school leavers have a minimum qualification of upper secondary school diploma, 71 per cent of US school leavers have a high-school diploma and 90 per cent of West German school leavers have a school leaving certificate. But although 84 per cent of UK school leavers have three or more CSE graded passes, the report's authors believe that these qualifications are of a far less demanding nature.

At the other end of the scale the United States and Japan lead in the numbers of graduates, with West Germany and the United Kingdom on more equal footing

Table 2 New degrees per million of population

	First degrees	Postgraduates
West Germany	1,200	260
Japan	3,250	170
UK	2,000	380
USA	4,100	1,700

(Table 2), although once again much depends on how comparable 'degree' qualifications are between countries. In terms of postgraduates, the United Kingdom moves into second place behind the United States, well ahead of both Japan and West Germany. An analysis on a subject basis shows the far heavier concentration of engineering graduates in both the US and Japan, but with Japan well behind the United Kingdom in terms of postgraduate scientists. What this brief synopsis suggests is that while the United Kingdom under-performs at lower educational levels, it scores far better at the much higher postgraduate levels. However, a recent report on British science, *Foresight in Science — Picking the Winners*, shows that despite this high output of scientists British science is not as successful as had been previously thought (*Nature* 312, 183; 1984).

Perhaps more revealing is the much higher priority that the three countries give to education and training, by the individual, employer and the state, and where the funding is regarded as necessary capital investment rather than as expenditure. Each however follows a different pattern. In both West Germany and Japan, employers

takes a strong lead and make a high financial investment in education and training and in recent years have been increasing this investment despite the recession. A particular feature of the United States is the much high investment by the individual in training and retraining and the diversity of, and flexibility, in training provision. The prime motivation in the United States is seen to be competitive/economic from the point of view of the individual and the employer, although the role of government, at federal and state level, is seen as more volatile.

The report's aim is to provide a stimulus for debate about how to improve VET, and hence competitiveness, in the United Kingdom. It comments that VET provision has to be geared to suit the country's specific circumstances and that there is no one single model of success. And it recognizes that while a direct relationship between VET and economic success cannot be proved, every body of opinion in the three countries believes that economic success cannot be possible without a strongly sustained VET effort. A large number of recommendations of how to improve the United Kingdom's position are presented. They include education and training for young people to cope with the demands of working life, with a goal of 85 per cent of school-leavers achieving acceptable standards in a range of core subjects — an aim which is now government policy. The recent UK initiatives such as YTS and TVEI are seen to be moves in the right direction. For employers, there is a need to increase commitment to both on-the-job and off-the-job training as an investment in the future. They also need to take a lead in identifying training priorities and in increasing the status and rewards that accrue to jobs in the business sector. Flexibility in terms of easier access to training and re-training are also needed, while government investment in human resources has to be maintained. Finally, individuals need to be more aware that their own skills are relevant to their employment roles. Education and training is a collaborative affair and needs to be given a higher priority by all concerned if competencies and competitiveness are to be improved. □

Table 1 Participation in education by 16–18 yr olds

Germany	86%
Japan	96%
UK	60%
USA*	73%

*Age range 16–24

Richard Pearson is at the Institute of Manpower Studies, Mantell Buildings, University of Sussex, Brighton BN1 9RF, UK.